



Los diseños de investigación cuantitativa en psicología y educación

Hugo Darío Echevarría

ISBN 978-987-688-166-1

e-book

Colección
Académico-Científica



Diseños de investigación cuantitativa en psicología y educación

Hugo Darío Echevarría



Universidad Nacional de Río Cuarto
Río Cuarto – Córdoba - Argentina

Echevarría, Hugo Darío

Diseños de investigación cuantitativa en psicología y educación / Hugo Darío Echevarría. - 1a ed.
- Río Cuarto : UniRío Editora, 2016.

Libro digital, PDF - (Académico científica)

Archivo Digital: descarga y online

ISBN 978-987-688-166-1

1. Metodología de la Investigación. I. Título.

CDD 001.42

Diseños de investigación cuantitativa en psicología y educación

Hugo Darío Echevarría

2016 © *Hugo Darío Echevarría*

2016 © *UniRío editora. Universidad Nacional de Río Cuarto*
Ruta Nacional 36 km 601 – (X5804) Río Cuarto – Argentina
Tel.: 54 (358) 467 6309 – Fax.: 54 (358) 468 0280
editorial@rec.unrc.edu.ar
www.unrc.edu.ar/unrc/comunicacion/editorial/

Primera edición: Marzo de 2016

ISBN 978-987-688-166-1



Este obra está bajo una Licencia Creative Commons Atribución 2.5 Argentina.

http://creativecommons.org/licenses/by/2.5/ar/deed.es_AR



Uni. Tres primeras letras de “Universidad”.

Uso popular muy nuestro; la Uni.

Universidad del latín “universitas” (personas dedicadas al ocio del saber),
se contextualiza para nosotros en nuestro anclaje territorial
y en la concepción de conocimientos y saberes contruidos
y compartidos socialmente.

El río. Celeste y Naranja. El agua y la arena de nuestro
Río Cuarto en constante confluencia y devenir.

La gota. El acento y el impacto visual: agua en un movimiento
de vuelo libre de un “nosotros”.
Conocimiento que circula y calma la sed.

Consejo Editorial

Facultad de Agronomía y Veterinaria

Prof. Laura Ugnia y Prof. Mercedes Ibañez

Facultad de Ciencias Humanas

Prof. Pablo Dema

Facultad de Ciencias Económicas

Prof. Ana Vianco y Prof. Gisela Barrionuevo

Facultad de Ingeniería

Prof. Jorge Vicario

Facultad de Ciencias Exactas, Físico-Quí-
micas y Naturales

Prof. Sandra Miskoski y Prof. Julio Barros

Biblioteca Central Juan Filloy

Bibl. Claudia Rodríguez y Bibl. Mónica Torreta

Secretaría Académica

Prof. Ana Vogliotti y Prof. José Di Marco

Equipo Editorial:

Secretaria Académica:

Ana Vogliotti

Director:

José Di Marco

Equipo:

José Luis Ammann, Daila Prado,

Maximiliano Brito, Ana Carolina Savino,

Daniel Ferniot

Índice

Prólogo.....	6
Capítulo 1. El proceso de investigación	9
Diseño en sentido amplio y restringido	10
La elección del tema, el relevo de los antecedentes y el enunciado del problema	14
Los objetivos de investigación	23
El marco teórico y las hipótesis de trabajo	30
Operacionalización de conceptos	32
Ejercicios: <i>Ejemplos de operacionalización</i>	34
Selección o construcción de un diseño de investigación.....	35
La población y la muestra.....	46
Los instrumentos de recolección de datos.....	48
El procesamiento de los datos.....	51
El cronograma de actividades	53
Los recursos disponibles y solicitados	54
Capítulo 2. Diseños rigurosos.....	57
La validez interna	57
La aleatorización en los diseños rigurosos	62
Diseño de dos grupos, con distribución al azar y postest solamente	64
Diseño de dos grupos, con distribución al azar, pretest y postest.....	65
Diseño de cuatro grupos de Solomon.....	67
Ejercicios	68
Capítulo 3. Diseños cuasiexperimentales	69
Diseños cuasiexperimentales sin grupo de control	69
Diseño de un grupo con postest únicamente	70
Diseño de un grupo, pretest y postest.....	72
Diseño de un grupo, doble pretest y postest	75
Diseño con pretest y postest con variable dependiente no equivalente*	75
Diseño con tratamiento removido*	76
Diseño de un grupo con tratamiento repetido*	77
Diseños de una serie cronológica.....	77
Investigación explicativa causal con un solo sujeto ($N=1$).....	78
Diseños cuasiexperimentales con grupo de control	82
Diseños sin pretest	82
Diseño de dos grupos y solo postest	82
Diseños con grupos duplicados y sólo postest*	83
Diseños de cohorte.....	83
Los diseños cuasiexperimentales de dos o más grupos, con pretest y postest.....	84
Diseños de dos o más series cronológicas*	85
Diseños compensados	86
Capítulo 4 Diseños no experimentales	89

Diseños no experimentales prospectivos	91
Diseño no experimental prospectivo de dos grupos, pretest y postest	92
Diseños retrospectivos.....	94
¿Grupos o situaciones? Diseño no experimental de una situación y sólo postest.....	96
Diseño no experimental de dos situaciones y sólo postest.....	97
Capítulo 5. Diseños factoriales	100
Diseños factoriales con control riguroso	102
Diseños factoriales cuasiexperimentales*.....	104
Validez interna y tipo de diseño.....	105
Capítulo 6. Investigación descriptiva.....	111
Los diseños transversales.....	112
Diseños longitudinales	114
Diseños de etapas emparejadas	116
La investigación descriptiva y el problema de la validez	117
Consideraciones finales	120
¿Qué diseño elegir?.....	120
¿Otros tipos de investigación?.....	123
Lógica en uso y lógica reconstruida	124
Apéndice A.....	127
Apéndice B. Ejemplos.....	128
Bibliografía citada	134

Prólogo

Existen varios libros destinados a tratar el tema del diseño de investigación y manuales que brindan una visión completa de la metodología, por lo que puede parecer innecesaria una nueva publicación. Sin embargo, existen algunos motivos que la justifican: en primer lugar, los distintos diseños se han formalizado con diferentes simbologías, incluso a veces esto ocurre en un mismo texto, lo que dificulta enormemente su comprensión e impide llegar a una visión integrada de los mismos, que permita al investigador decidir cuál puede ser la mejor opción para su proyecto.

En segundo lugar, un aspecto central como el problema de la validez, no se trata en forma sistemática para algunas de las opciones disponibles: apenas se han realizado apreciaciones aisladas para los diseños descriptivos y factoriales.

En tercer lugar, puesto que el texto también puede ser de interés para estudiantes de grado, a quienes suele resultarles difícil vincular lo relativo a los diseños de investigación con otros elementos que intervienen durante todo el proceso, siendo muy común que confundan algunos de ellos, se torna necesario desarrollar los primeros en relación con el último.

Por ello, este libro tiene los siguientes objetivos:

1. Dar una visión global del proceso de investigación, en la que pueda apreciarse el lugar que ocupa el diseño de investigación.
2. Desarrollar una simbología que pueda aplicarse tanto a los diseños experimentales como a aquellos que no lo son, y presentarlos a todos con la misma forma de representarlos.
3. Ofrecer una primera aproximación al problema de la validez tal como esto fue realizado por Campbell y Stanley (2005), con el aporte más reciente de Shadish et al. (2002).
4. Analizar las ventajas y limitaciones de cada uno y los requisitos que exige su aplicación.

En el Capítulo 1, presento una visión global del proceso y el proyecto de investigación, ubicando el diseño en relación con ambos. En los capítulos sucesivos, trataré las distintas alternativas, comenzando por los estudios experimentales en sentido estricto (Capítulo 2). Después presentaré los cuasiexperimentales (Capítulo 3), los no experimentales (Capítulo 4), los factoriales (Capítulo 5) y los descriptivos (Capítulo 6).

En relación a la diagramación de los diseños de investigación cualitativa, en otro lugar (Echevarría, 2011), he realizado una sistematización de ellos partiendo de la concepción de Yin (2009), representándolos con la misma simbología que la usada para los diseños de investigación cuantitativa y, por este motivo, no los incluyo aquí. En las consideraciones finales, haré algunas reflexiones sobre el modo de usar los distintos diseños disponibles.

Este libro es una versión corregida y ampliada de los diseños cuantitativos presentados en Echevarría, H. 2005. *Los diseños de investigación y su implementación en Educación*. HomoSapiens Ediciones, Rosario.

Se inició en el marco del proyecto *Creatividad en los procesos de investigación de los alumnos universitarios* aprobado y subsidiado por la Secretaría de Ciencia y Técnica de la Universidad Nacional de Río Cuarto, sin cuyo estímulo hubiera sido imposible finalizarlo.

También quiero expresar mi reconocimiento a aquellos que de distinta manera colaboraron en la concreción de este trabajo:

En relación a la primera versión, a Karina Bergé y Eduardo Soler, quienes realizaron valiosas indicaciones que, por un lado, facilitan la lectura y, por otro lado, permitieron un mejor tratamiento de los conceptos involucrados. A Gloria Vadori, por las sugerencias que permitieron corregir detalles de redacción en una de las versiones avanzadas del escrito.

Respecto de esta versión, a Mabel Grillo y Viviana Machiarolla, quienes de distintas formas me ayudaron a concretarlo.

A aquellos alumnos que leyeron algunos borradores y que con mucho entusiasmo hicieron aportes para mejorar la legibilidad del texto.

A mi esposa y a mi madre, que incondicionalmente me apoyaron; a mi hija Lucrecia aún sin estar entre nosotros, y a mi hijo Agustín; porque sin quererlo y sin saberlo, fueron dos luces que brillaron en la oscuridad iluminando mi camino en algunos momentos difíciles que tuve que atravesar.

Capítulo 1

El proceso de investigación

Es muy común afirmar que la investigación científica se desarrolla a través de un proceso que, como tal, está compuesto por una serie de etapas, aunque existen diferencias en el modo de nombrarlas, el número que contiene y el orden en que se llevan a cabo. En la investigación cuantitativa, generalmente se piensa en un proceso lineal, pues estas etapas se cumplirían en un orden determinado¹.

En este trabajo, considero que estas grandes etapas son básicamente cuatro: a) la planificación, que consiste básicamente en la elaboración del proyecto de investigación, b) el trabajo empírico (o trabajo de campo), c) el procesamiento de los datos y, d) la comunicación científica (que incluye presentaciones orales e informes escritos). Cada una de estas actividades, a su vez, incluye a otras, como se aprecia en el Esquema 1.1 (ver más adelante). A su vez, todas están planteadas como actividad,

¹ El trabajo científico por su complejidad no siempre avanza según un orden estrictamente lógico y predeterminado. El conocimiento se construye con muchos avances y retrocesos, y a veces se comienza por aquello que, según la recomendación de los metodólogos, debe concretarse en otros estadios de la investigación. No obstante, esto puede traer innumerables inconvenientes, incluso a veces insalvables, por lo que se debe seguir lo más fielmente posible un orden lógico, ya que esto nos brinda mayores garantías de cumplir con la finalidad para la que se inició el trabajo. Si en lugar de etapas, hablamos de tipos de actividades tal vez tengamos una idea más realista del trabajo científico, aunque en la metodología cuantitativa es fuerte la idea de etapas que se desarrollan en un orden predeterminado.

pero algunas de ellas también pueden verse como componentes del proyecto e incluso del informe de investigación.

Esto puede llevar a confusiones, pues se nombra del mismo modo a la actividad y el resultado de la misma. Por ejemplo, para relevar los antecedentes y enunciar el problema, el investigador deberá concretar una serie de tareas: buscar bibliografía, resumirla o ficharla, escribir una síntesis en la que se muestren los avances en relación al tema elegido, y finalmente enunciar la pregunta o problema de investigación; lo que se traduce en un trozo de texto, que es mucho más reducido que todo lo resumido o escrito en una primera instancia. Lo importante es advertir, que ese trozo de texto, que es el resultado de la actividad *relevar los antecedentes y plantear el problema*, se llama casi del mismo modo que ella. Así, *algunas* de las actividades que se muestran en el Esquema 1.1, también pueden verse como partes o elementos del proyecto o del informe de investigación (aunque no todas integran ambos documentos). A su vez, en el Cuadro 1.1 se aprecian las preguntas que se deben responder en cada actividad, preguntas *que no deben confundirse con las preguntas o problema de investigación*.

En lo que resta del Capítulo 1, en primer lugar me referiré al sentido en que se usará el término diseño en este trabajo, luego trataré en forma muy resumida los aspectos involucrados en el proyecto de investigación, para continuar (desde el Capítulo 2 en adelante), con los diseños de investigación o tipos de estudios que podemos desarrollar en psicología y educación.

Diseño en sentido amplio y restringido

El modo en que se define el término *diseño* resulta controvertido. Mientras que algunos le dan *un alcance sumamente general, que implica prácticamente la totalidad de los aspectos que involucra la investigación, otros lo conciben en forma muy restringida*².

² Indudablemente que esta dicotomía es una simplificación, pues existen algunas posiciones intermedias, no pudiéndose incluir en estos extremos.

Esquema 1.1. El proceso de investigación

Planificación (El proyecto de investigación)	-> Trabajo empírico o de campo	-> Procesamiento, reducción y presentación de los datos	-> Comunicación científica
Elección de un tema y revisión bibliográfica	Prueba de los instrumentos de recolección de datos	Creación de variables	Corrección de los aspectos teóricos y metodológicos
Planteo del problema o preguntas de investigación	Corrección de los instrumentos de recolección de datos	Ingreso de los datos	Análisis de los datos
Formulación de objetivos y justificación de la investigación	Entrada al campo	Control de errores	Interpretación y discusión de los resultados
Explicitación del marco teórico, hipótesis de trabajo y operacionalización de conceptos	Recolección de los datos, experimentación	Reducción y presentación de los datos	
Delimitación de la población de interés y el modo de selección de la muestra			
Construcción de un diseño de investigación			
Elaboración de los instrumentos de recolección de datos			
Selección de las técnicas de análisis de los datos			
Construcción del cronograma de actividades y descripción de los recursos disponibles y solicitados			

Cuadro 1.1. Preguntas que se contestan en cada tarea de investigación y apartados en que podrían incluirse

Apartados posibles para proyectos o informes	Elección de un tema	¿Qué área o tópico me interesa, puedo o debo hacer la investigación?
Introducción Antecedentes, planteo del problema y objetivos	Revisión bibliográfica	¿Qué se investigó?, ¿sobre qué hay acuerdo?
	Planteo del problema o preguntas de investigación	¿Qué dudas identifico?, ¿con cuál/es quiero trabajar?
	Formulación de objetivos	¿Qué me propongo aportar para solucionar el problema?
	Metas	¿Cuánto me propongo lograr de cada objetivo?
	Justificación de la investigación	¿Por qué es importante lograr esos objetivos?
Marco teórico	Explicitación del marco teórico	¿Que supongo?, ¿cómo defino los principales términos que usaré?
	Hipótesis de trabajo	¿Qué enunciados pongo a prueba? (en estudios explicativos)
	Operacionalización de conceptos	¿Qué hipótesis auxiliares y definiciones debo incluir para inferir los indicadores?, ¿utilizaré algún índice?, ¿qué relaciones establezco entre indicadores y categorías o valores?
Metodología	Delimitación de la población de interés	¿A quienes pretendo generalizar los resultados? (si pretendo generalizar), ¿para quienes pretendo mostrar descripciones o explicaciones plausibles?
	Tamaño de la muestra y cantidad de casos a seleccionar	¿Cuántos casos indagaré?, ¿cómo los seleccionaré?
	Construcción de un diseño de investigación	¿Formaré grupos para comparar, de qué forma?, ¿cuando aplicaré los instrumentos de recolección de datos?, ¿aplicaré tratamientos, en qué momentos?
	Elaboración o selección de los instrumentos de recolección de datos	¿De qué modo mediré u observaré las variables?
	Selección de las técnicas de reducción de los datos	¿Cómo reduciré los datos? ¿Qué estadísticas calcularé?
Cronograma de actividades	Construcción del cronograma de actividades	¿Qué actividades llevaré a cabo, en qué momentos?
Recursos disponibles y solicitados	Descripción de los recursos disponibles	¿Qué recursos tengo para implementar el proyecto?
	Descripción de los recursos solicitados	¿Qué recursos solicito para implementar el proyecto?

En este trabajo, utilizo el término *diseño de investigación* en el segundo sentido (*restringido*), es decir, para referirme a *la estrategia o plan usado en el momento empírico (trabajo de campo), que comprende la recolección de datos, aunque puede incluir otras acciones, como por ejemplo, la introducción de algún tratamiento experimental o el proceso de negociación con los actores sociales en la investigación cualitativa.*

Para construirlo, debemos tomar decisiones relacionadas con los siguientes puntos:

1. Cantidad de grupos a formar.
2. Cantidad de sujetos en cada grupo y modo de distribuirlos.
3. Observaciones o mediciones que se realizarán y momento en que las mismas se llevarán a cabo.
4. Variaciones o tratamientos que se introducirán.

Tampoco debemos confundir lo que es el diseño con el proyecto, el informe o el proceso de investigación. Éste último se refiere a todas las acciones que lleva a cabo el científico (Samaja, 2004), entre las cuales la planificación (elaborar el proyecto) forma el primer grupo de ellas, mientras que la comunicación científica se traduce en uno o más informes, que pueden ser básicamente orales o escritos. El diseño (en sentido restringido), que es una parte del proyecto, debe estar también descrito en el informe y nos orienta en el momento empírico de la investigación. Como dije más arriba, a algunos elementos que conforman el proyecto, el informe y el proceso los llamamos del mismo modo (por ejemplo, el enunciado del problema o los objetivos de investigación), pero si hablamos del proceso deben entenderse como la actividad que conduce a ellos, mientras que si nos referimos al proyecto o al informe, son el resultado de esta actividad.

En relación a la *comunicación científica*, Yuni y Urbano la definen como “las tareas de sistematización y de organización discursiva que permitan exponer con la mayor claridad y de acuerdo a las convenciones propias del campo científico los hallazgos de investigación” (2003: 38). Esta expresión (*comunicación científica*), tiene la ventaja sobre el término *informe*, de no hacer alusión exclusivamente a un único in-

forme, por escrito y cuando ya finalizó el trabajo empírico. Dentro de ella podemos incluir, por ejemplo, las presentaciones en congresos (que suelen ser orales exclusivamente o combinando esta modalidad con la escrita), que no necesariamente se dan luego de haber recolectado todos los datos, pues el investigador puede contar con resultados parciales o reflexiones teóricas que considera pertinente presentar antes de completar el trabajo de campo³.

En algunos contextos acotados puede ser adecuado el término *informe final*, (como por ejemplo en la realización de un trabajo monográfico para cumplir con las exigencias del dictado de una materia), pero en términos generales es preferible hablar de *comunicación científica*.

Veremos los principales diseños que podemos usar en la investigación cuantitativa, pero antes me referiré a todos los elementos que implica el proyecto y el lugar que ocupan en relación con éste, aunque algunos de ellos serán tratados también como actividad.

La elección del tema, el relevo de los antecedentes y el enunciado del problema

Podemos decir que los *antecedentes* están conformados por todos los documentos a los que tenemos acceso y son previos a nuestra investigación. Son en algún sentido, lo que la antecede. Algunos utilizan la expresión *estado del arte* para referirse a ellos, pero lo más importante, es advertir que el término *antecedentes* se utiliza con dos sentidos diferentes. La definición que acabo de dar corresponde a uno de ellos, ya que, *en un sentido muy general*, representan todo lo que se hizo antes de que nosotros iniciáramos nuestro proyecto. Pero en un *sentido restringido*, incluyen únicamente los trabajos empíricos, y su análisis es fundamental para establecer hasta dónde se avanzó en un área determinada. Por ello, debemos comenzar por hacer una revisión bibliográfica en la que podemos hallar datos, opiniones, conceptos, sugerencias, sistemas o reglas para intervenir en relación a un problema social determinado, hipótesis, supuestos, definiciones, descripción de instrumentos de recolección de datos, técnicas y programas para su análisis, planteos

³ Aquí notamos que a veces las etapas no se dan según el orden lógico que se plantean en los textos de metodología.

metodológicos o epistemológicos, etc. Dependiendo de lo que el investigador tiene en mente, cada uno de estos adquirirá mayor o menor importancia.

Algunas preguntas pueden orientarnos en el relevo de los antecedentes, incluyo a continuación algunas de ellas: ¿que se investigó en el tema x?, ¿qué descripciones del fenómeno existen?, ¿son satisfactorias?, ¿dejan dudas?, ¿qué dudas dejan?, ¿qué teorías se propone para explicar estos hechos?, ¿se han contrastado?, si esto se hizo, ¿los distintos estudios coinciden en los resultados hallados o son contradictorios?, ¿se hicieron estudios que permiten interpretar las acciones sociales consideradas?, ¿son creíbles?

Actualmente el relevamiento bibliográfico se hace mediante el uso de bases de datos informáticas con programas diseñados específicamente para esta tarea o al menos con aquellos buscadores de propósitos generales como el conocido Google, aunque si buscamos antecedentes con los productos de esta empresa, Google Académico (Google Scholar) representa la mejor opción. Se diseñó para hallar textos científicos exclusivamente y si bien no tenemos disponibles algunas revistas especializadas, tiene la enorme ventaja de ser gratuito, con acceso libre a una gran cantidad de documentos de muy buena calidad. A veces nos brinda únicamente la información de determinada publicación sobre un tema dado (autor, título del artículo, nombre de la revista, etc.) pero no tenemos acceso al informe completo, pero con un poco de paciencia obtenemos una cantidad de documentos suficientes para hacer nuestro trabajo.

Existen varias bases de datos de este tipo, aunque algunas de acceso restringido y el material se encuentra en diferentes idiomas, que generalmente es el inglés⁴. Una base de datos de uso libre sumamente interesante para quienes el castellano es la lengua materna es Redalyc (Red de Revistas Científicas de América Latina y el Caribe, España y Portugal; Sistema de Información Científica, Universidad Autónoma del Estado de México, cuya página principal es <http://www.redalyc.org/>). Esta base de datos, además de contar con una gran cantidad de publicaciones de buena calidad, permite acceder a muchos documentos completos en forma totalmente libre, tiene un buscador sumamente sencillo y potente, y cuenta con la posibilidad de hacer búsquedas avan-

4 En el momento de escribir estas páginas se cuenta con la posibilidad de acceder a ellas desde los servidores de universidades e instutos públicos de algunos países latinoamericanos.

zadas filtrando los documentos hallados sobre la base de varios criterios simultáneamente.

Si el investigador no tiene conocimiento de los términos claves para la búsqueda bibliográfica, puede hacer algunas preguntas y organizarla sobre la base de ellas. Por ejemplo, supongamos que está interesado en la motivación de los estudiantes de la escuela primaria, podría preguntarse: ¿hay estudios que describen la motivación de los alumnos en este nivel?, si no están motivados, ¿qué hipótesis se indagaron para explicarlo?, ¿en qué variables influye la motivación?, ¿se hicieron experiencias para revertir el problema?, ¿se plantearon propuestas sin llevarse a la práctica?

Las dudas que dejan los estudios previos pueden deberse entre otras razones, a la inexistencia de datos empíricos que describen el fenómeno que queremos estudiar, o que den sustento a ciertas hipótesis explicativas relacionadas al mismo, aunque también podría suceder que existan pero que estos datos sean contradictorios. Si hallamos estudios cualitativos, al relevar la bibliografía debemos analizar si las interpretaciones del fenómeno halladas son convincentes, aceptadas por los investigadores y cuáles son las posibles líneas para continuar con estos trabajos. En cierta manera, lo fundamental en este momento de la investigación, es *establecer una línea lo más clara posible entre lo que es conocido satisfactoriamente y aquello que deja dudas*.

Por supuesto que establecer esta línea es extremadamente complejo, pues requiere de mucho conocimiento sobre el tema abordado, pero precisamente las dudas que podemos identificar nos indican problemas de investigación posibles para trabajar en ellos. Retomaré el tema de los antecedentes más adelante, pues para seguir avanzando se torna necesario definir lo que son los problemas de investigación y diferenciarlos de los problemas sociales.

Sabino (1996) distingue el *problema práctico* del *problema de investigación*, lo que me parece sumamente conveniente, pues es muy común que se produzca una confusión entre ambos. El primero se refiere a cualquier dificultad que se nos presenta en la vida cotidiana; se resuelve realizando alguna acción e implica un proyecto de intervención, en el cual, la principal finalidad es modificar la realidad. Cuando es elevada la cantidad de sujetos que tienen un *problema práctico* determinado,

éste se transforma en un *problema social*, esto es, *que la sociedad percibe como tal*.

El *problema de investigación* se da cuando estamos ante algo desconocido o mal conocido, resolverlo implica obtener un nuevo saber (Sabino, 1996) y remite a un proyecto de investigación. Éste puede ayudar, e incluso a veces resultar imprescindible para resolver un problema social⁵. También es posible que un proyecto de intervención (social) incluya como subproyecto a otro de investigación, o que este último dé origen a aquél, pero son diferentes, pues los de intervención tienen por objetivo modificar la realidad, mientras que los de investigación, la obtención de un nuevo conocimiento⁶.

Ander-Egg y Aguilar Idáñez (2004: 16) mencionan los proyectos sociales y culturales, que están destinados a “producir determinados bienes y servicios capaces de satisfacer necesidades o resolver problemas”. Es decir, con ellos se interviene sobre la realidad para solucionar algún problema social. Por eso, podemos considerar sinónimos a los términos proyecto social y proyecto de intervención.

En lo que sigue, me referiré exclusivamente a los proyectos de investigación, en los cuales lo primero que debemos hacer es elegir un tema o área de interés. A veces, esta última nos viene dada por una necesidad que proviene de algún problema social (por ejemplo, somos contratados para conocer la situación escolar de los niños que trabajan en la calle); otras veces, podemos elegirlo libremente. En el último caso, la decisión puede estar influida, por ejemplo, por algún hecho que despertó nuestra curiosidad, por la sugerencia de algún colega o compañero, o simplemente por una cuestión de preferencia personal. De cualquier manera, es el primer paso que debemos dar, pues *los problemas científicos se plantean a partir de los antecedentes que existen sobre el tema* y, sin elegir el último, difícilmente se puedan buscar los primeros.

Además, para plantear el problema, primero debemos contar con una perspectiva amplia del estado de avance del conocimiento en el área seleccionada, observando en la bibliografía aquellos aspectos que

5 Desde el punto de vista metodológico, un problema social se da cuando una o más variables están fuera del rango de los valores aceptados o deseables, como por ejemplo, elevada tasa de mortalidad infantil en un determinado sector geográfico o país.

6 En la investigación acción esta distinción se diluye, pues este tipo de trabajos persigue ambas cosas al mismo tiempo, aunque puede haber predominio de uno u otro objetivo.

se repiten y sobre los que hay acuerdo entre los investigadores, diferenciándolos de aquellos que no hay acuerdo, pues estos últimos indican *problemas de investigación* (dudas que los antecedentes dejan y con las que se quiere trabajar). Para esta finalidad, puede ser de utilidad construir una tabla como la que se presenta en el Apéndice A. En ella podemos ubicar en las filas todos los elementos de cada autor leído según se señala en las columnas. Así, si la examinamos en dirección a las últimas, tendremos las coincidencias en los distintos autores para cada aspecto y, si miramos en dirección a las filas, los elementos que incluye cada texto. Por ejemplo, en la sexta columna, podemos apreciar quiénes plantean un determinado problema social y, si hacemos lo propio con la quinta, los datos que ofrecen para fundamentar la existencia del mismo. En las celdas podemos insertar párrafos textuales del material leído, aunque también otros que no tengan esta característica, incluso algunos que sean muy resumidos o esquemáticos. Es conveniente, en todos los casos, registrar la página en que se halla cada una de las ideas o enunciados que se insertaron, pues seguramente con algunos se necesitará volver a leer el original.

En lugar de una tabla como la sugerida, algunos investigadores prefieren realizar fichas; otros, cuadros sinópticos o esquemas, pero cualquiera sea la opción elegida, *lo importante es que llegue a una visión resumida del estado del conocimiento en el momento en que intenta plantear el problema o las preguntas de investigación.*

Cuando tenemos esta visión global y esquemática del estado del arte en el área seleccionada, estamos en condiciones de intentar la formulación del problema, que es una de las etapas iniciales, y también cruciales, del proceso de investigación. Padua (1987: 28) sostiene que “todo problema de investigación comienza siempre como un problema de teorización” y Van Dalen y Meyer (1971: 145) nos dan nuevas pistas para entender lo que representa este momento del trabajo científico:

“Algún hecho confunde o provoca molestias a alguien; una inquietud corroe su tranquilidad de espíritu, hasta que después de determinar con precisión qué es lo que lo perturba, encuentra algún medio para resolverlo (...) *El problema se materializa cuando el estudioso percibe que algo no está bien o requiere una explicación más profunda*” (cursiva agregada).

Al hablar del *problema de investigación*, nos referimos siempre a *ambigüedades, vacíos o contradicciones en el conocimiento de un área deter-*

minada. Si ocurre lo primero, estamos ante un problema conceptual que no necesariamente da origen a un trabajo empírico⁷. Como dije más arriba, en este libro me referiré únicamente a las investigaciones empíricas.

Cuando la dificultad está dada por vacíos (o desconocimiento), hay ciertos aspectos de la realidad que se ignoran parcial o totalmente. Problemas de este tipo podrían ser los siguientes: se desconoce si los alumnos de primer año de la ciudad *X* leen comprensivamente, o no se sabe por qué algunos alumnos, a pesar de estar muy motivados y poseer buen nivel intelectual, fracasan en la escuela.

Por último, el problema puede originarse en contradicciones en el conocimiento existente y en este caso tenemos distintas alternativas: a.1) dos (o más) teorías dan explicaciones diferentes de un mismo fenómeno, a.2) se observan contradicciones entre proposiciones pertenecientes a una misma teoría, b) entre una teoría y ciertos datos (hechos), c.1) existe evidencia empírica a favor de una teoría, pero también de tipo refutatoria (contradicción entre algunos datos que corroboran una teoría y otros que la refutan), c.2) se dan contradicciones entre datos descriptivos⁸. La Tabla 1.1 sintetiza los distintos tipos de problemas de investigación que podemos considerar.

También existen discrepancias relativas al modo de formular el problema, pues para algunos puede asumir la forma de una pregunta, pero para otros de un enunciado descriptivo. Quizás la posición más aceptada sea la que sostiene que debe hacerse del primer modo. Kerlinger, (1990: 18) es uno de los autores que sostiene este punto de vista aunque piensa exclusivamente en estudios explicativos: “Un *problema*, entonces, es una oración o aseveración interrogativa en la cual se pregunta: ¿qué relación existe entre dos o más variables?” (cursivas del autor).

7 Entiendo por *investigación empírica* a aquella que comprende a los estudios experimentales, los no experimentales, los descriptivos, los cualitativos y los mixtos. Es opuesta, por un lado, a los trabajos de clarificación conceptual o de revisión bibliográfica y, por otro lado, a los referidos a las ciencias formales.

8 En este caso se pueden plantear estudios que intentan replicar investigaciones ya realizadas.

Tabla 1.1. Tipos de problemas de investigación

Existencia de ambigüedades	Existe conocimiento, pero no es suficientemente preciso. La definición de algunos términos no es clara o no están bien operacionalizados. La relación entre algunos términos al enunciar las hipótesis o entre hipótesis no es precisa. Estos problemas dan origen a investigaciones no empíricas (que no se tratan en este libro).
Desconocimiento	No se investigó sobre ciertos aspectos o sectores de la realidad, la evidencia empírica es insuficiente para corroborar o refutar ciertas hipótesis
Contradicciones	Existe conocimiento y es preciso, pero contradictorio. Las contradicciones pueden ser: a.1) entre teorías; a.2) dentro de una teoría; b) entre datos y teoría; c) entre datos (algunos datos corroboran una teoría y otros la refutan o entre datos de distintos estudios descriptivos).

Van Dalen y Meyer son más amplios, en tanto admiten que también el problema se puede formular a través de un enunciado descriptivo:

“Un enunciado completo del problema incluye todos los hechos, relaciones y explicaciones que se revelaron como importantes en el análisis. No basta con enumerar estos factores o incluirlos en una lista, sino que es necesario poner de manifiesto las relaciones que existen entre ellos. Debe encuadrárselos *en un enunciado descriptivo o en una pregunta* que indique con claridad qué información ha de obtener el investigador para resolver el problema” (Van Dalen y Meyer, 1971: 157, cursiva agregada).

Pese a ser una idea difundida, *no es correcto afirmar que el problema de investigación es una pregunta, pues un interrogante no refleja en sí mismo una dificultad. En todo caso, el problema está dado por el hecho de que dicha pregunta no ha sido satisfactoriamente respondida.*

Por esta razón creo que es más pertinente decir que se puede enunciar el problema o las preguntas de investigación e, incluso, ambas cosas. Lo importante es que identifiquemos los huecos, vacíos o contradicciones en el conocimiento existente, pues esto es fundamental para avanzar sobre terreno firme en las siguientes actividades.

Si enunciamos las preguntas de investigación, éstas pueden referirse a una relación entre dos o más variables (como vimos en la cita de Kerlinger). Por ejemplo, podrían ser del siguiente tipo: “¿la motivación influye en el rendimiento?”, “¿el rendimiento y la motivación están correlacionadas?” Sin embargo, también tenemos otras posibilidades que no

se limitan a plantear supuestas relaciones entre variables. Por ejemplo, “¿es elevado el porcentaje de niños hiperactivos en el nivel primario de la ciudad de Río Cuarto?”, “¿aumentó el fracaso escolar en la provincia de Córdoba en el año 2016 con relación a 2010?”

Si no se realizan preguntas de investigación, la formulación sintética del problema, que *debe desprenderse del análisis de los trabajos previos*, puede asumir la forma de un enunciado descriptivo, por ejemplo, del siguiente tipo: “No se encontraron trabajos que analizan si *el tipo de actividades académicas* influye en la *motivación*”, “son pocos los estudios que analizan la relación entre el *tipo de actividades académicas* y la *motivación*”, “no hay estudios que den cuenta del porcentaje de niños hiperactivos en el nivel primario de la ciudad X”, “no se conoce si aumentó el fracaso escolar en la provincia de Córdoba en el año 2016 con relación a 2010”. Todos estos ejemplos se refieren a *casos de desconocimiento o ignorancia acerca de ciertos aspectos de la realidad social*.

Pero vimos que también el problema puede estar dado por contradicciones en el conocimiento existente, como los que asumen las formas siguientes: “En algunos trabajos se sostiene que *el tipo de actividades académicas* influye sobre en la *motivación* y en otros que esto no sucede”, “algunos autores encontraron una correlación alta entre *el tipo de actividades académicas* y la *motivación* y otros baja”, “se hallaron estudios que muestran que el fracaso escolar en el nivel primario de la ciudad X es importante, en otros resulta insignificante”, “según A el fracaso escolar en la provincia de Córdoba en el año 2016 fue del tanto por ciento, para B es más alto (o más bajo)”.

Estamos aquí mostrando la forma sintética que asume el enunciado del problema o las preguntas de investigación, pero cualquiera sea la opción que elegimos, *ésta es el resultado de un proceso de análisis que es mucho más amplio que ella*. Algunos de estos elementos deben incluirse en esta formulación. Así, retomando el ejemplo que recién mencionamos, si afirmamos que “algunos trabajos sostienen que X influye sobre Y y otros que esto no sucede”, *deben citarse y comentarse brevemente estos estudios*, incluyendo las correlaciones que se hallaron, si éstas resultaron significativas, y en general *todos aquellos elementos que permitan fundamentar que de la bibliografía existente se pueden inferir las afirmaciones contradictorias incluidas en la formulación del problema*.

Cuando enunciamos el problema no nos estamos refiriendo a la realidad de la que hablan los investigadores, sino al conocimiento que tenemos sobre ella. Veamos algunos ejemplos de enunciado de problemas de investigación.

“La mayor parte de la investigación sobre las técnicas grupales se refiere a aplicaciones en la escolaridad primaria y secundaria. *No abundan los estudios en el nivel universitario.* Una excepción es una reciente experiencia innovadora realizada en la Facultad de Ciencias Políticas y Sociología de la Universidad Complutense de Madrid (España) (**Carballo**, 2002), que perfila un nuevo modelo de enseñanza, aunque difícilmente pueda encuadrársela como una investigación en sentido estricto” (Roselli, 2008, **negrita en el original**, cursiva agregada).

Decir que “no abundan los estudios en el nivel universitario” es lo mismo que afirmar no se ha investigado suficientemente en esta área. Este caso se relaciona a los vacíos en el conocimiento mencionados más arriba; en la práctica difícilmente hallemos áreas no investigadas, sino que, por lo general los problemas se plantean porque la evidencia recabada es insuficiente, lo que puede suceder tanto en estudios explicativos como descriptivos.

Veamos un ejemplo del último tipo:

“En una investigación reciente en la que se utilizaron escalas de puntuación para padres y maestros como método de identificación del TDAH⁹, se calculó la prevalencia del TDAH para niños marabinos en un 7,19% [7]. Se administraron estas escalas a padres y maestros de niños entre 6 y 12 años que asistían a escuelas regulares. Sin embargo, *la identificación mediante las escalas de puntuación no tuvo confirmación clínica*” (Montiel-Nava et al., 2003, Pp.: 815, cursiva agregada).

En este último ejemplo, los autores presentan un problema de contradicción entre datos descriptivos, ya que éstos no resultaron coincidentes cuando se aplicaron escalas estandarizadas y se realizaron diagnósticos clínicos del TDAH. Plantean claramente un problema de contradicción entre datos descriptivos, lo que hace que se justifique realizar nuevas indagaciones.

⁹ Attention deficit hyperactivity disorder.

Blázquez (2006) nos ofrece otro ejemplo de problema de investigación, pero se refiere a un tema poco investigado. Al hablar de los bailes de cuarteto en la ciudad de Córdoba (Argentina), sostuvo:

“Si bien existen algunas investigaciones sobre una danza y unas sonoridades originadas a mediados del siglo XX en la región central de Argentina llamada Cuarteto o Cuarteto Cordobés, las mismas se basan preferentemente en la figura del artista más famoso a nivel nacional -Carlos ‘La Mona’ Giménez- o en los aspectos psicológicos. Estos trabajos se asemejan, a pesar de sus diferencias metodológicas, *por prestar escasa atención* a la performance de la danza, a los procesos de subjetivación de los agentes involucrados y desatender o ignorar el género sexual como dimensión analítica.

Este silencio resulta bastante significativo más aún cuando el *Cuarteto* se presenta explícitamente como un tipo de danza social que organiza intencionalmente el encuentro (hetero)sexual y para lo cual divide estrictamente (no sólo) a través de la coreografía a los sujetos en hombres y mujeres” (Blázquez, 2006: 97).

Decir que algunos trabajos se parecen en que *prestan escasa atención* a un tópico determinado, es equivalente a afirmar que este tema no se investigó adecuadamente, por lo que estamos ante un auténtico problema de investigación. Pero además de estas afirmaciones, se podrían plantear también las preguntas de investigación lo que no siempre resulta redundante, como es el caso de este informe. El autor, luego lo hace, lo que le permite precisar en mayor medida el problema (véase en el siguiente apartado la cita que presento de Blázquez al analizar los objetivos).

Los objetivos de investigación

Entendemos que el problema siempre supone dificultades, ya sea por ignorancia o contradicciones en el conocimiento existente. Por lo tanto, *su formulación debe explicitar en la mayor medida posible estas dificultades y justificarlas con la bibliografía*. A su vez, de él se deben desprender los objetivos, en el sentido de que ellos se orientan a realizar algún aporte para solucionarlo, aunque no se trata de resolverlo en un solo trabajo y para siempre.

Los objetivos más fundamentales de la investigación social pueden ser de cuatro grandes tipos:

- Describir en profundidad o comprender casos particulares.
- Desarrollar (construir) teoría.
- Describir cuantitativamente una muestra o una población.
- Contrastar una o más hipótesis explicativas.

Los dos primeros son los objetivos que *por lo general* se persiguen con los estudios cualitativos (que como ya dije no trato aquí por haberle dedicado otro trabajo), el tercero, con los descriptivos (cuantitativos) y el cuarto, con los explicativos. A su vez, el investigador puede combinar más de un tipo en un estudio determinado, aunque aquí, *por razones didácticas, la exposición se realizará partiendo del supuesto de que su proyecto considerará objetivos de una sola clase.*

Por otro lado, cuando éstos son demasiado amplios para poder concretarlos, se torna necesario diferenciar entre objetivos generales y específicos. En este caso, se debe cuidar que los últimos estén incluidos en algunos de los primeros, pues suele ser frecuente que se enuncien objetivos específicos sin vinculación con los generales o que tengan el mismo grado de amplitud que éstos. En todo caso, siempre deberá tenerse presente que esta distinción debe servir sobre todo para ayudar al investigador a desglosar un aspecto muy amplio en otros más acotados, con la finalidad de poder abordarlo. Por ejemplo, el siguiente objetivo es demasiado amplio: Conocer si la forma de estudio influye en la deserción estudiantil.

Por el modo en que está enunciado, no resulta orientador de las acciones que se deben seguir desarrollando. La situación cambia si formulamos otros más específicos como los siguientes, que están incluidos en el primero y son más orientadores que él:

- Conocer si el tiempo dedicado al estudio influye en la deserción,
- Indagar si la carencia de estrategias de estudio adecuadas favorece la deserción.

Estos dos últimos objetivos, entre otras cosas, nos dicen qué marco conceptual puede ser usado, y a partir de él qué instrumentos de recolección de datos podrían seleccionarse o construirse.

Según Yuni y Urbano (2003) los objetivos tienen los siguientes componentes: verbo, variable, unidades de análisis/observación y contextualización. Al primero lo consideran la “acción cognitiva”, las variables son los aspectos, atributos, aspectos o cualidades que observaremos o mediremos en los “sujetos o entidades” acerca de los cuales construiremos algún conocimiento nuevo y el contexto es más difícil de definir, pues por un lado incluye las condiciones bajo las cuales se producirá el conocimiento, pero también aspectos generales como por ejemplo, el tiempo en que se realizará la recolección de datos, la región geográfica, los tipos de sujetos o entidades que abarcará, etc. Veamos algunos ejemplos de los informes recién citados en relación al problema de investigación.

“El objetivo de este trabajo es confirmar las cifras de prevalencia del TDAH en una comunidad venezolana y describir las características clínicas y demográficas de los sujetos identificados” (Montiel-Nava et al., 2003, Pp.: 815).

Así, los autores plantean hacer una nueva descripción con la finalidad de cotejarla luego con las existentes hasta ese momento. Roselli (2008: 90) en cambio, se propuso otra cosa. “El objetivo de la investigación fue, como se ha dicho, comparar dos modelos alternativos de organización de la enseñanza: individual y grupal”.

Blazquez lo hace de un modo más detallado, pero también con claridad.

“En el presente trabajo exploramos las formas de clasificación de las diferentes experiencias y modos de transitar la femineidad de acuerdo al discurso de las propias adolescentes que asisten a los *bailes de Cuarteto*. Según nuestra observación, los estereotipos utilizados por y para la clasificación de las mujeres (con)funden distinciones de clase, género, raza y orientación sexual. Estas formas purificadoras están organizadas según dos paradigmas de diferenciación, uno racial/estético/moral y otro sexual/genérico/erótico, articulado en un término que cuando es aplicado a un sujeto particular lo coloca a éste en una posición alejada y subordinada respecto al sujeto de la enunciación que se construye como *normal*. Siguiendo a Butler (2002[1993]), podemos entender que en esta producción performativa, citacional

o iterativa del sujeto *normal* se constituye simultáneamente en el espacio de lo abyecto habitado por aquellos cuerpos que no importan y cuya presencia amenaza y confirma las fronteras que separa a unos de otros. ¿Quiénes son las mujeres que están (pre) destinadas a ocupar esas posiciones degradadas o degeneradas en los mundos del *Cuarteto cordobés*? ¿Cómo se materializan los cuerpos *normales*? ¿Cuáles son sus fantasmas? ¿Qué formas de vigilancia corporal se instauran para evitar los deslizamientos hacia el peligroso camino de lo abyecto? ¿Cómo se regulan las prácticas identificatorias? ¿Cómo se articulan, bajo el régimen de heterosexualidad obligatoria (Rich, 1999[1980]) que imperan en los *bailes*, estos paradigmas de diferenciación en la producción de los seres abyectos? Estas son algunas de las preguntas, construidas entre las noches en el baile y las tardes entre los libros, que orientan el presente trabajo” (Blazquez, 2006: 98-99).

En el párrafo de Blazquez se presentan tres elementos. En primer lugar el objetivo, luego (en la siguiente oración) una síntesis de los resultados hallados, lo que queda claro pues sostiene que según su observación los estereotipos son de determinada forma, luego una cita de Butler que es un supuesto de amplia generalidad y le sirve para interpretar lo hallado y finalmente las preguntas de investigación, lo que le permite precisar en mayor medida el problema planteado más arriba (que presenté al analizar el problema).

Veamos otros ejemplos de enunciados de objetivos de investigación. Gil Cuadra et al. (2003: 27), lo hizo del siguiente modo: “La contribución de este trabajo consiste en describir y caracterizar las concepciones y creencias que sobre enseñanza aprendizaje de las matemáticas mantienen los profesores de secundaria andaluces”; mientras que Chirinos (1999: 50) se propuso “Identificar el nivel de conocimiento y las actitudes sobre sexualidad y las prácticas sexuales de las estudiantes adolescentes en colegios secundarios de Lima, Perú”.

La Tabla 1.2 muestra los distintos componentes de los objetivos de investigación de los trabajos citados. En este caso no incluí el contexto, ya que no siempre los investigadores lo explicitan inequívocamente.

Tabla 1.2. Componentes de los objetivos en algunas investigaciones

<i>Publicación</i>	<i>Verbo</i>	<i>Variable o aspecto indagado</i>	<i>Unidades de análisis</i>
(Montiel-Nava et al., 2003).	Confirmar	Cantidad de niños con TDAH	Una comunidad venezolana, “escolares entre 3 y 13 años que asistían a escuelas regulares en el Municipio Maracaibo del Estado Zulia, en Venezuela”
Roselli (2008)	Comparar	Sistemas de organización de la enseñanza	Alumnos de nivel superior
(Blazquez, 2006: 98-99)	Explorar	“Formas de clasificación de las diferentes experiencias y modos de transitar la femineidad”	“Adolescentes que asisten a los <i>bailes de Cuarteto</i> ” en la ciudad de Córdoba, Argentina
Gil Cuadra et al. (2003)	Describir, caracterizar	Concepciones y creencias sobre enseñanza y aprendizaje de las matemáticas	Profesores de secundaria andaluces
Chirinos (1999)	Identificar	Nivel de conocimiento, actitudes sobre sexualidad, prácticas sexuales	Estudiantes adolescentes en colegios secundarios de Lima, Perú

Hay que notar que los objetivos de estudios de contrastación de hipótesis tienen una estructura diferente, aunque a primera vista parecen muy similares¹⁰. Sus componentes son los siguientes (para una hipótesis con dos variables): Verbo, *si*, (variable independiente), influye en, (Variable dependiente). *En la Tabla 1.3 se presentan algunos ejemplos.*

10 La estructura planteada por Yuni y Urbano mencionada más arriba es válida para los estudios descriptivos y cualitativos, para los de contrastación de hipótesis es demasiado general y no permite advertir todos sus elementos. Por otro lado, también debe tenerse en cuenta que lo planteado aquí es válido para estudios que contrantan hipótesis que incluyen sólo dos variables.

Tabla 1.3. Componentes de los objetivos explicativos y ejemplos

<i>Verbo</i>	<i>si</i>	<i>variable independiente</i>	<i>influye en</i>	<i>variable dependiente</i>
Conocer	si	el nivel educacional de los padres	influye en	la deserción
Indagar	si	la carencia de estrategias de estudio	influye en	la deserción

Sin embargo, debemos tener presente que a veces se enuncian con una estructura diferente aunque es equivalente a la mostrada la Tabla 1.3. Por ejemplo, podríamos expresarlos del siguiente modo:

En este trabajo se ponen a prueba las siguientes hipótesis:

- el nivel educacional de los padres influye en la deserción,
- la carencia de estrategias de estudio adecuadas favorece la deserción.

Otro punto que suele generar conflictos es la relación entre objetivos generales y específicos, en ese sentido ésta puede ser de diferentes tipos, esto es, los objetivos específicos pueden ser, respecto de uno general: a) Una parte, b) Una especificación o meta y, c) Pasos para llegar a él¹¹.

Veamos ejemplos de objetivos específicos según cada forma de concebirllos para el siguiente objetivo general:

Indagar la relación entre la comprensión lectora y el rendimiento en alumnos de primer año del colegio.

a) Conocer la correlación entre comprensión de la microestructura y rendimiento, conocer la correlación entre comprensión de la macroestructura y rendimiento (según un marco conceptual la comprensión de la microestructura y la macroestructura son parte de la comprensión lectora, por ello estos objetivos forman parte del general).

b) Calcular el coeficiente de correlación entre comprensión lectora y rendimiento en al menos el 90% de la muestra seleccionada.

c) Desarrollar un instrumento para conocer la comprensión lectora de los alumnos, seleccionar una muestra de alumnos, aplicarle los ins-

¹¹ Tal vez no todos aceptarían los puntos b) y c).

trumentos de recolección de datos para medir la comprensión lectora, analizar las propiedades psicométricas de los instrumentos de recolección de datos desarrollados para medir la comprensión lectora, obtener la nota de los parciales, condición final y de exámenes finales de los alumnos, estudiar la forma de la distribución del rendimiento, calcular la correlación entre comprensión lectora y rendimiento y la diferencia de medias en rendimiento entre los alumnos de alta y baja comprensión lectora.

A veces el enunciado de las hipótesis precede al de los objetivos, pues éstos y la construcción del marco teórico forman un todo que sólo desglosamos con fines analíticos. Pero, ¿qué relación existe entre el problema y las hipótesis de investigación¹²?

Van Dalen y Meyer (1971) sostienen que la hipótesis es la solución del problema. Siguiendo a Kerlinger, podemos decir que “una *hipótesis* es una afirmación en forma de conjetura de las relaciones entre dos o más variables” (1990: 19, cursiva en el original).

Por ejemplo, de la forma “ X influye sobre Y ”, o “ X e Y están correlacionadas”.

A su vez, los objetivos del estudio podrían ser poner a prueba una o más hipótesis, y en este caso los llamaré *estudios de contrastación de hipótesis o estudios explicativos*. La última forma suele usarse, pues una hipótesis que tiene un cierto grado de sustento empírico puede ser usada para dar explicaciones a ciertos fenómenos o hechos.

También la finalidad podría ser *describir* algún sector de la realidad. Por ejemplo, el siguiente objetivo es descriptivo y remite a un estudio de esta clase: “conocer el porcentaje de fracaso escolar en el nivel primario de la ciudad de Río Cuarto desde el año 1990 al 2015”.

Tanto los estudios descriptivos como explicativos, suponen un marco teórico en el cual, entre otras cosas, deben definirse las variables mencionadas en los objetivos. En este trabajo diferencio entre supuesto e hipótesis, tal como lo planteo a continuación.

12 Por razones didácticas, aquí asumo que el investigador abordará un solo problema, formulará un solo objetivo y si su estudio es explicativo, considerará una sola hipótesis entre todas las que podría elegir.

El marco teórico y las hipótesis de trabajo

Los *supuestos* que conforman las teorías son enunciados generales⁽¹³⁾ y si algunos de estos enunciados se ponen a prueba, entonces en esa investigación lo consideramos *hipótesis de trabajo*, *hipótesis de investigación* o simplemente *hipótesis*, lo que implica que la diferencia entre ambos está dada por la función o el rol que cumplen en cada caso. Si se ponen a prueba son las *hipótesis*, si los consideramos verdaderos sin cuestionarlos, son *supuestos*. Un mismo enunciado puede ser supuesto en una investigación e hipótesis en otra.

Las hipótesis de trabajo pueden ser de correlación (o asociación) y causales. Las *hipótesis de correlación* postulan la existencia de una concomitancia o covariación entre dos variables o la existencia de una ocurrencia conjunta de ciertas categorías para cada individuo de la población. Por ejemplo, si decimos que el género se asocia a la elección vocacional, estamos postulando una relación entre la categoría “femenino” y cierta clase de carreras, e inversamente, entre “masculino” y otro tipo de carreras diferentes. Las *hipótesis causales* afirman una relación de causalidad entre variables, e intervienen al menos una variable independiente (causa) y al menos una variable dependiente (efecto¹⁴).

Todo problema se ve desde un marco teórico y toda hipótesis está enmarcada en una *teoría*, que le da sentido a los términos que contiene. Podemos definir a la última como un conjunto de supuestos articulados lógicamente entre sí. Algunos autores conciben como sinónimos a los términos “supuesto” e “hipótesis”. Como ya vimos, considero que los *supuestos* son todas las afirmaciones que no se contrastan en la investigación, es decir, se los acepta como verdaderos (podemos decir que se relegan al conocimiento de fondo no problemático). Esto no supone que sean incuestionables, sino que a los efectos de poder avanzar, tomamos la decisión metodológica de considerarlos verdaderos. Los diferencia-

13 En cambio, los datos son enunciados particulares que describen un hecho o una síntesis de ellos.

14 A veces se incluyen las hipótesis descriptivas dentro de las hipótesis de trabajo, lo que no hago, debido a que por lo general las primeras no se consideran tales, sino que se las ve como la información empírica. Por ejemplo, si afirmo que el puntaje promedio en una prueba de matemáticas fue de 67, esta afirmación puede verse como una hipótesis descriptiva o un dato que surgió de la realidad. Podría postularse a priori que este grupo obtendrá tal puntaje (o que superará cierta medida) y en este caso sería la hipótesis de trabajo. Pero también podría inferirse el valor una vez que se recabó la información, con lo cual vemos a este enunciado como una síntesis de la información recolectada. Esto último es lo que por lo general hacen los investigadores.

mos de las *hipótesis*, pues éstas son las que se contrastan con la realidad, esto es, las ponemos a prueba en nuestro trabajo. El *marco conceptual*, por otro lado, está formado por los conceptos y supuestos que son relevantes por ejemplo, para plantear el problema, construir los instrumentos de recolección de datos e interpretarlos una vez que contamos con ellos¹⁵, lo que significa que es más restringido que la teoría, pero a su vez, el *modelo* es más limitado que aquél:

“El modelo en la ciencia es una serie de proposiciones coordinadas y subordinadas que más frecuentemente se expresan mediante fórmulas que describen o explican relaciones reales de objetos. Los modelos son esquemas (...). Las teorías se prueban mediante modelos. Un modelo suele ser sólo una parte de la teoría. Los modelos, pues, buscan facilitar la contrastación de teorías” (Castañeda Jiménez, 1996: 163-164).

Podemos decir que el *modelo* está formado por las hipótesis cuando se han operacionalizado sus conceptos y precisado sus relaciones. Esto nos lleva al siguiente tema, la operacionalización de conceptos, que está vinculada fundamentalmente al desarrollo teórico, aunque a veces se la considera una parte de la construcción de los instrumentos de recolección de datos. Pero antes veamos brevemente lo que son las definiciones y las diferencias que existen con los datos empíricos¹⁶.

Las definiciones establecen una relación de equivalencia entre lo que definen y lo definido. En cambio, los datos son enunciados particulares que describen un hecho o una síntesis de ellos. Por ejemplo:

“Apego inseguro-resistente o ansioso-ambivalente: reacciona fuertemente a la separación. Cuando la madre regresa buscan el consuelo y el reencuentro pero pueden mostrar rabia o pasividad: no se calman con facilidad, tienden a llorar de una manera desconsolada y no retoman la actividad de exploración” (Carreiras, et al., 2008: 110).

En este caso, el autor separa los dos componentes con los dos puntos (“:”), pero a veces se usa la palabra *es*, u otras como *se entiende, considero*,

15 Se suelen considerar sinónimos los términos “teoría” y “marco conceptual”, aquí sostengo que el último sólo incluye la parte de la primera que es relevante para nuestro estudio. Esto significa que un supuesto podría formar parte del marco conceptual de un estudio y estar excluido de otro, sin embargo, ambos pertenecen siempre a la misma teoría.

16 Desde la epistemología se ha problematizado lo que son los datos y las teorías que suponen. Dada la naturaleza introductoria de este libro adopto una posición ingenua al respecto, como así también con otros conceptos, como por ejemplo los de hipótesis, supuesto, etc.

se da, etc. Por ejemplo, el autor podría haber dicho, “El *apego inseguro-resistente o ansioso-ambivalente* se da cuando el sujeto reacciona fuertemente a la separación y cuando la madre regresa busca el consuelo y el reencuentro pero puede mostrar rabia o pasividad. Además, no se calma con facilidad, tiende a llorar de una manera desconsolada y no retoman la actividad de exploración”.

Veamos otra definición:

“Creencias: las verdades personales indiscutibles sustentadas por cada uno, derivadas de la experiencia o de la fantasía, que tienen un fuerte componente evaluativo y afectivo (Pajares, 1992). Las creencias se manifiestan a través de declaraciones verbales o de acciones (justificándolas)” (Gil Cuadra et al. 2003: 28).

El autor podría haber dicho, por creencia entiendo las verdades personales... O en este trabajo se considera que las creencias son las verdades personales, etc.

Operacionalización de conceptos

La discusión acerca de la función, momento, posibilidad, necesidad o conveniencia de la operacionalización de conceptos es muy antigua dentro de las ciencias sociales. Ella involucra aspectos puramente filosóficos vinculados a la validez y fundamentación de la ciencia, pero también otros de tipo técnico que son relevantes a la hora de planificar y concretar una investigación. En relación con el primer punto de vista, hallamos dos extremos, entre los cuales se ubican una gran cantidad de posiciones intermedias. Por un lado, desde la tradición cuantitativa, están aquellos que creen en la necesidad de una operacionalización completa y previa al trabajo empírico, aclarando en forma precisa el significado de cada concepto utilizado (ver por ejemplo, la caracterización de este paradigma en Popkewitz, 1988). Por otro lado, desde los enfoques cualitativos, se ubican quienes no creen en tal posibilidad o simplemente en su inconveniencia¹⁷.

17 Esta posición no es aceptada por todos, pues para algunos la vinculación entre lo teórico y lo empírico se debe plantear en la investigación cualitativa en los mismos términos que en la cuantitativa (véase cita de Samaja en la próxima nota al pie).

En este apartado, trataré sintéticamente la operacionalización de conceptos en el plano técnico, tal como se plantea desde el modelo cuantitativo. Este recorte se realiza simplemente con la finalidad de acotar el trabajo, aunque también puede orientar posteriores discusiones, ya sea entre los partidarios del punto de vista estándar o para confrontarlo con quienes sostienen el enfoque cualitativo¹⁸.

Sin desconocer la complejidad de la discusión sobre este tema y adoptando una posición ingenua, comenzaré por definir lo que entiendo por esta tarea, para establecer los pasos que involucra y sus características específicas.

Podemos decir que *operacionalizar* un concepto teórico es establecer sus vinculaciones con la realidad, con lo cual lo transformamos en un concepto empírico¹⁹. Este proceso involucra ciertas etapas, que se presentan de un modo secuencial, aunque en la práctica es posible que se cambie el orden de las mismas, como así también que se superpongan algunas de ellas. Me basaré en Cea D'Ancona (1999)²⁰, aunque agregando un paso que considera implícitamente.

En la *primera etapa*, realizamos la definición del concepto, desde una teoría que en este momento puede tener distintos grados de explicitación y sistematización. En un *segundo momento* y derivándose del anterior, establecemos las distintas dimensiones (subvariables), que no son otra cosa que los grandes aspectos que involucra o componen el concepto.

Una vez establecidas las dimensiones (suponiendo que no tuviéramos nuevas subdimensiones derivadas de las primeras) -*tercera etapa*-

18 Por ejemplo, sería interesante analizar si en las investigaciones cualitativas podemos hablar de dimensiones, indicadores, categorías (o valores) e índices como lo hago aquí para la cuantitativa (a lo que habría que agregar la matriz de datos, tema que no trato pero que está supuesto en estos conceptos). Una idea más restringida sostiene Samaja (1994), cuando afirma que “todos los datos de toda investigación científica se pueden analizar en cuatro elementos básicos, que guardan entre sí relaciones estructurales: unidad de análisis, variables, valores e indicadores” (: 376). Es más restringida por cuanto no incluye a los índices en el análisis de los datos cualitativos.

19 Esta definición es incorrecta si adoptamos por ejemplo, una posición constructivista, pues en ella, la realidad es una construcción y como tal se traduce en enunciados. En cierto sentido esta definición es una simplificación desarrollada con fines puramente didácticos, sólo aceptable para un realismo compatible con concepciones positivistas ortodoxas.

20 Esta autora se basó en un clásico trabajo de Lasarsfeld; tomé el de ella, pues los términos con que menciona a las etapas resultan más claros que los que usa aquél.

debemos establecer las categorías o valores de cada una de ellas²¹. Estas categorías son el conjunto de posibilidades que puede asumir una dimensión (o subvariable, aunque esta definición también es válida para las variables, pues desde el punto de vista lógico son idénticas, sólo difieren en el lugar que ocupan en la definición operacional). En una *cuarta etapa*, establecemos los indicadores de cada categoría de las subvariables y finalmente *-quinta etapa-* realizamos un *índice*, que nos permite construir la variable, al sintetizar en ella todos los valores o categorías que asumen las dimensiones. En otros términos, el índice nos permite asignarle un valor único a la variable a partir de los valores atribuidos a las dimensiones, y puede traducirse en una fórmula matemática, o en un conjunto de relaciones cualitativas entre las categorías de las dimensiones y las de la variable a observar.

Ejercicios: Ejemplos de operacionalización

Una forma de medir la popularidad

Tome dos personajes

Ponga sus nombres en un buscador de documentos Web (por ejemplo el conocido Google), de a uno por vez.

Cuente las páginas en que aparece cada uno.

Esto ya da una estimación de la popularidad de los personajes, pero tiene un gran inconveniente. Se darán muchos falsos positivos y negativos. Para evitar lo primero se pueden agregar los pasos siguientes:

Seleccione una muestra al azar de las páginas de cada uno de los personajes. Por ejemplo, por muestro sistemático tal como se explica este procedimiento en el apartado correspondiente.

Entre a cada una de las páginas seleccionadas.

Verifique si se trata de un hallazgo correcto o de un falso positivo.

Obtenga el porcentaje de hallazgos correctos.

Corrija la cifra original usando regla de tres simple.

Calcule su propio índice de inflación

Tome al menos 10 productos que Ud. consume habitualmente.

El primer día de un mes cualquiera, vaya al supermercado en el que habitualmente realiza las compras y anote cuánto cuesta cada uno.

²¹ Ésta es la etapa que agregué y que no consideran los autores mencionados.

El primer día del mes siguiente, haga lo mismo (puede tomar más días y luego corregir lo hallado calculando lo que hubiera sucedido si hubiera tomado un mes).

Calcule cuánto aumentó cada producto en porcentaje.

Obtenga un promedio de todos los porcentajes.

Reflexione: ¿es válido ese índice de inflación?, ¿qué puntos fuertes y débiles tiene?

Si obtiene valores muy distintos a los de los organismos oficiales, ¿a qué puede deberse?

¿Cuáles son las dimensiones o aspectos que considera (por ejemplo alimentos, educación, etc.)?

Selección o construcción de un diseño de investigación

¿Qué es un diseño de investigación? Comencemos por ver un ejemplo de un estudio experimental (ver recuadro).

Vidal-Abarca, (2000) analizó la relación que existe entre ciertas características de los textos y la cantidad de proposiciones recordadas, de ideas erróneas incluidas en los protocolos de recuerdo y de preguntas inferenciales que los alumnos contestaron luego de su lectura. Partiendo de un texto original (O), elaboró tres versiones alternativas (T; I y T+I). En la versión T, se agregaron proposiciones que estaban implícitas en el texto original, la versión I incorporó expresiones nuevas con la finalidad de favorecer la conexión entre las ideas previas y el contenido del texto y finalmente, la versión T+I agregó ambos tipos de proposiciones. El autor describe el procedimiento seguido en los siguientes términos:

“participaron 64 chicos y chicas de 13-14 años que estaban estudiando el tema «la revolución rusa». Los chicos se distribuyeron aleatoriamente en cuatro grupos leyendo cada uno bien el texto original tomado de un libro de texto real (Ciencias Sociales de 8º, Editorial Anaya, 1987), bien una de las tres versiones alternativas. (...) El procedimiento de aplica-

ción de las pruebas fue el siguiente. En una primera sesión, los chicos: a) contestaron un test de conocimiento previo sobre el tema, b) estudiaron la versión que les correspondió durante 18 minutos, y tras una breve tarea distractora c) escribieron lo que recordaban del texto. La segunda sesión tuvo lugar una semana más tarde. En ellas los chicos: a) escribieron de nuevo lo que recordaban del texto, y b) respondieron a una prueba formada por siete preguntas inferenciales pudiendo consultar a voluntad el pasaje que habían leído” (p. 149).

Entre los resultados que mostró el estudio, se halló que en las pruebas de recuerdo inmediato y diferido, solamente con la versión T se observó un recuerdo significativamente superior al texto original aunque esta diferencia se debió fundamentalmente al recuerdo de ideas de bajo nivel y se registró la mayor cantidad de proposiciones erróneas en ella.

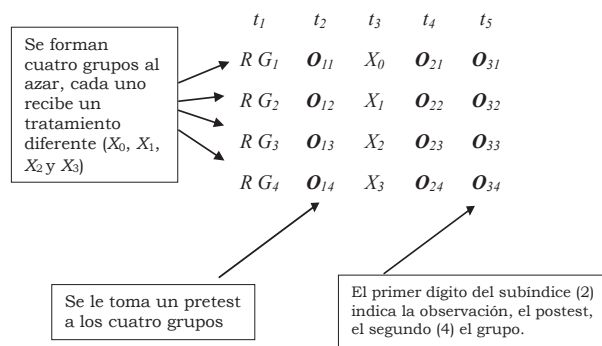
¿Qué hizo Vidal-Abarca (2000) en la parte empírica de su trabajo? Podemos establecer los pasos que siguió durante el mismo, partiendo del momento en que ya había seleccionando una muestra:

- 1) Conformó cuatro grupos asignando los sujetos a ellos al azar.
- 2) A todos los sujetos les tomó una prueba sobre el tema *la revolución rusa*.
- 3) A cada grupo le asignó una condición diferente, es decir, le hizo leer el texto original (O), o bien, las versiones alternativas T, I y T+I, sobre el mismo tema.
- 4) Luego de leer los textos, le pidió a los niños que escribieran lo que recordaban del mismo (en forma inmediata a la lectura).
- 5) Una semana más tarde, le solicitó que escribieran nuevamente lo que recordaban del texto y que respondieran a siete preguntas inferenciales sobre lo que habían leído.

Podemos ahora esquematizar los elementos del diseño, tratando que cada aspecto esencial del mismo quede representado por un símbolo, como en el Diagrama 1.1. En él, t_1 , t_2 , t_3 , t_4 y t_5 , indican los momentos o tiempos en que ocurre cada etapa, son los pasos 1) a 5) en la enume-

ración que hice más arriba. En el primero (t_1), se conformaron cuatro grupos *al azar* (G_1 , llamado grupo de control y G_2 a G_4 , los grupos experimentales), lo que está señalado por la R . Esta letra proviene de la palabra inglesa “random”, que significa precisamente “azar” (más adelante explicaré con más detalle lo que significa la asignación al azar y porqué ella es tan importante en la investigación experimental). Aquí se le llama grupo de control al conformado por los niños que se usaron lo habitual y experimentales a las modificaciones que introdujo el experimentador. A veces el grupo de control recibe un placebo o no recibe tratamiento alguno²². También se le puede llamar tratamientos a todos ellos, ya que el placebo o no recibir nada pueden verse como tales. En este último caso podemos decir que estamos ante el tratamiento nulo.

Diagrama 1.1. Diseño de cuatro grupos, con distribución al azar, pretest y posttest inmediato y diferido



Este diseño permite considerar una variable independiente que cuenta con cuatro valores o categorías (X_0 , X_1 , X_2 y X_3). No debe confundirse con el de cuatro grupos de Solomon. Con este último, en dos grupos no se realizan pretest y usa una variable independiente con sólo dos valores (*ver capítulo siguiente*).

En cada columna, se halla todo lo que coincide en el tiempo, mientras que lo que se ubica en una fila determinada, se da con el mismo grupo. La “O” indica alguna observación o medición que se realiza a todo el grupo, sintetizada en algún estadístico que lo representa (como la media o

²² El placebo es un tratamiento o una situación que para los sujetos resulta similar o igual al tratamiento verdadero, pero que se sabe que no produce el efecto estudiado sobre la variable dependiente. Por ejemplo, inyecciones de agua destilada en un experimento médico que trata de probar la influencia sobre una enfermedad de una droga que se suministra de esta forma. Se realiza para evitar el efecto psicológico que tiene el hecho de que el sujeto recibe un tratamiento que cree que es eficaz para superar o reducir su problema.

el porcentaje de respuestas correctas, por ejemplo). Si se efectúa luego del tratamiento o los tratamientos, las llamamos *posttest*. X_1 a X_3 denotan la existencia de los tratamientos experimentales (los textos alternativos que leyeron los niños) y X_0 la ausencia de él o lo que se cumple habitualmente. En este caso la lectura del texto original.

Si por ejemplo, alguien que no leyó el resumen del artículo de Vidal-Abarca (2000) que está más arriba (ni tampoco el trabajo completo), ve el diagrama del diseño, puede advertir varias cosas sobre el mismo: el *número de grupos que se formaron* (cuatro en este caso), la *forma en que se asignaron los sujetos* a estos grupos (al azar, lo que queda simbolizado por la letra *R*), los *momentos en que se aplicaron los instrumentos de recolección de datos* (las *O*) para medir la variable dependiente (comprensión lectora) y por último, la *cantidad de tratamientos que se aplicaron y en qué grupos* (las *X*, que representan las categorías de la variable independiente). Estos son los elementos que incluye el diseño en la concepción restringida del mismo mencionada más arriba²³.

Nótese que los subíndices de las observaciones (*O*) en realidad debieran ir separados de una coma, por ejemplo el primero debiera ser *1,1* y no *11*, ya que no es el número *11* lo que queremos representar, sino el primer elemento de dos listas. La primera lista es la de las observaciones, que como se ve en el diagrama, coincide en las columnas. Es decir, si miramos las observaciones en sentido vertical, notamos que todas tienen el mismo subíndice (*1* la primera, *2* la segunda y así sucesivamente). Mientras que el segundo subíndice, coincide en los grupos y por lo tanto, es el mismo para toda una fila (puede verse el número *1* para todo el primer grupo, el *2* para el segundo, etc.).

Este diagrama nos ofrece la forma, esto es, la estructura del diseño, independientemente de su contenido. Introduciendo otros contenidos en esa estructura, podemos generar otra investigación. Por ejemplo, queremos indagar si la enseñanza grupal es más adecuada que la individual en relación con un tema determinado, como podría ser el concepto de *zona de desarrollo próximo* en alumnos universitarios. Entonces conformamos cuatro grupos al azar, le tomamos una prueba a todos los

23 En Echevarría (2011) utilicé un concepto más amplio de diseño, aunque sigue siendo restringido. En él incluí las posibles variaciones en los instrumentos de recolección de datos propios de la investigación cualitativa, como así también los momentos de su análisis. Esta distinción no es necesaria si tratamos sólo con diseños cuantitativos que suponen una cierta constancia o equivalencia en el proceso de recolección de datos y su evaluación.

sujetos sobre el tema, luego, un grupo aprende individualmente, otro lo hace en grupo de 2 sujetos, el tercero en grupo de 4 y el último de 6. Después de llevar a cabo la actividad de enseñanza, le aplicamos a todos los sujetos un cuestionario sobre el mismo tema, lo que repetimos una semana más tarde. En este caso, hemos hecho otra investigación con el mismo diseño.

Notemos que, para el caso en que la variable independiente tiene sólo dos categorías o valores, X_0 también puede representar un tratamiento opuesto a X_1 o simplemente un tratamiento diferente. Además, en t_1 no se observa, sino que solamente se conforman los grupos. Aunque puede parecer más natural comenzar con t_1 al iniciar las observaciones, reservé este momento para la conformación de los grupos o la decisión acerca de quienes se va a observar, pues es un aspecto central del diseño que se vincula directamente a su validez.

Así, contamos con una forma de formalizar los diseños, y analizarlos de un modo más abstracto y general. Existen infinitas investigaciones que podrían hacerse con cada uno de los diseños, de modo que ciertas reflexiones que haremos sobre las ventajas y limitaciones de uno determinado en algunas de ellas, nos servirán para analizar otras, incluso, aunque aún no se hayan concretado (estamos en la fase de diseñarla).

Otra manera de definir el *diseño*, equivalente a la anterior es la siguiente: es la forma que tiene la estrategia seguida en la fase empírica o trabajo de campo. Si se trata de un estudio en que se siguió un proceso lineal, el diseño también puede verse como los pasos a cumplir en el mismo. Pero como no siempre esto sucede (sobre todo en la investigación cualitativa) podemos hablar de un modo general de estrategia.

En una situación concreta, *el investigador primero debe preguntarse qué diseños puede aplicar (en función del contexto en que se desarrollará el trabajo), luego entre los posibles, cuál es mejor, es decir, con cuál tiene menos probabilidades de cometer errores de inferencias, lo que no es otra cosa que hacer consideraciones sobre validez de su estudio*. Dada la importancia que tienen para el tema que trato, en el capítulo siguiente retomaré los elementos del diseño y la forma de simbolizarlo. Por ahora, simplemente reitero las dos ideas básicas que orientan este libro: *por un lado mostrar distintas alternativas con que cuenta el investigador a la hora de plantear sus proyectos (sin pretender ser exhaustivo) representadas con la misma simbología, y realizar una introducción al tema de la validez*.

El diseño del trabajo de Vidal-Abarca (2000) que acabamos de ver, es de tipo experimental, pero, puesto que este término se usa con sentidos diferentes, veamos sintéticamente algunos de éstos. En primer lugar, se lo utiliza para mencionar al *método* que adoptan las ciencias llamadas fácticas y que supone algún tipo de observación o experimentación con la realidad, por oposición a aquellos casos en que esto no ocurre (como en las ciencias formales que usan el método axiomático). En segundo lugar, en referencia a los *estudios experimentales* con dos sentidos: incluyendo a todos los que requieren manipulación de variables en investigaciones explicativas, o exclusivamente a aquellos que consideran más de un grupo y en los cuales los sujetos se distribuyen en ellos al azar.

En este trabajo, cuando hable de *estrategia, estudio o diseño experimental*, me referiré a aquellos en que se manipulan una o más variables. Cuando el investigador realiza un *experimento riguroso o un cuasi experimento*, cualquiera sea el grado de control que éste posea, *modifica algún aspecto de la realidad (variable independiente) para conocer qué efecto tiene sobre algún otro (variable dependiente)*.

No existe una posición unánimemente aceptada sobre el modo de clasificar los distintos tipos de estudios en la investigación social. Aquí consideraré que *existen tres grandes clases de diseños: los cualitativos, los cuantitativos y los mixtos*. Esta clasificación toma como principal criterio el carácter de los datos que se usarán (ver el Cuadro 1.2.1), agrupando por un lado a todos los que suponen el uso de instrumentos de recolección y análisis de datos cualitativos, por otro lado a los que pretenden describir o explicar la realidad con datos cuantitativos y, finalmente, los *mixtos*, que se caracterizan por combinar estas posibilidades. Esta definición del enfoque cualitativo no es compartida por todos los especialistas, ya que por lo general se la caracteriza por adoptar un enfoque interpretativo. Adhiero a la idea de Erickson (1989) quien diferencia entre investigación cualitativa con una orientación positivista y otra con orientación interpretativa, aunque por todas las aristas que tiene, no puedo profundizar en este tema aquí.

Cuadro 1.2.1. Diseños de investigación en Ciencias Sociales

Diseños de investigación	Cuantitativos
	Cualitativos
	Mixtos

Por razones didácticas, los diseños mixtos no se abordarán en este trabajo. Un tratamiento exhaustivo puede verse en Tashakkori y Teddlie (2010), Creswell (2003) y Teddlie y Tashakkori (2009). Tampoco lo haré con los estudios de psicogénesis, que pueden ser de tipo cualitativo o mixto, para los cuales puede verse el Capítulo 3 de Echevarría (2005) y a los diseños cualitativos les dediqué otro texto (Echevarría, 2011a).

A su vez, respecto de los diseños cuantitativos, consideraré, por un lado, a los de contrastación de hipótesis y, por otro lado, los descriptivos, estableciendo subclases dentro de cada uno (Cuadro 1.2.2). A los primeros luego los clasificaré por el grado de control que tiene el investigador sobre las condiciones en que se realiza el estudio, estableciendo una clase con aquellos en que se manipula la o las variables independientes (las modifica para indagar su efecto sobre otra u otras) y los grupos se forman al azar (diseños rigurosos, en el Capítulo 2 veremos porque es tan importante este requisito); otra clase en la que el investigador logra sólo lo primero (cuasiexperimentales) y una tercera en que no se dan ninguna de las dos condiciones. Para los estudios descriptivos tomaré una caracterización muy utilizada que considera tres tipos: transversales, longitudinales y de cohortes (que también se los conoce como de etapas emparejadas). El Cuadro 1.2.3 muestra estas mismas opciones con otras subclases y se recomienda volver a él antes y luego de leer cada capítulo.

Cuadro 1.2.2. Clasificación general de los diseños cuantitativos

Contrastación de hipótesis causales (Explicativos)	Estrictos o rigurosos	
	Cuasiexperimentales	Sin grupo de control
		Con grupo de control, sin pretest
		Con grupo de control y pretest
		Compensados
	No experimentales	
Descriptivos	Transversales	
	Longitudinales	
	Etapas emparejadas	

Cuadro 1.2.3. Clasificación detallada de los diseños de investigación cuantitativa (Plan de libro)

Contrastación de hipótesis causales (Explicativos)	Estrictos o rigurosos	
	Cuasiexper. sin grupo de control	Sólo posttest
		Pretest y posttest
		Doble pretest y posttest
		Pretest, posttest, variable depend. no equivalente
		Con tratamiento removido
		Con tratamiento repetido
		De una serie cronológica
		N=1
		De dos grupos y sólo posttest
	Cuasiexper. con grupo de control, sin pretest	Grupos duplicados y sólo posttest
		De cohorte
	Cuasiexper. con grupo de control y pretest	Con pretest y posttest
		Dos o más series cronológicas
	Cuasiexperimentales Compensados	
	No exper.	Prospectivos
		Retrospectivos
	Factoriales	
Descriptivos	Transversales	
	Longitudinales	
	Etapas emparejadas	

Exper.: experimentales; Depend.: dependiente.

En el difundido trabajo de Campbell y Stanley (2005) desarrollaron la primera gran sistematización de los diseños de investigación refiriéndose a los de contrastación de hipótesis causales (en 1966 se publicó la primera edición en inglés, la que utilizo aquí es la novena reimpresión en español), pero la publicación del libro de Shadish et al. (2002) marca una ruptura importante con éste, no obstante, en algunos aspectos, el primero nos sigue resultando de utilidad, por lo que en muchas ocasiones tomaré algunos de sus conceptos.

Por otro lado, se debe tener presente que en una misma investigación podemos incluir tanto objetivos cualitativos, como descriptivos o de contrastación de hipótesis. Si esto ocurre, podríamos considerar que estamos haciendo más de un estudio al mismo tiempo o que cabe la posibilidad de opciones intermedias entre las que se mencionan aquí.

En todo caso, el lector deberá tener presente que la clasificación que se presenta pretende ser una guía que lo oriente sobre las alternativas posibles y que, se pueden incluir objetivos de distinto tipo en un mismo trabajo, para lo cual se deberán implementar distintos diseños, en el sentido acotado con que se usa el término diseño en este libro.

En una primera aproximación, podemos decir que con la *investigación descriptiva*, simplemente nos muestra cómo es la realidad, lo que existe y en qué proporciones, o se propone conocer qué valores asumen ciertos parámetros para una población determinada y cómo varían con el tiempo.

Cuando realizamos un estudio de contrastación de hipótesis, ponemos a prueba al menos una hipótesis que establece una relación de causalidad entre dos o más variables, lo que podemos hacer manipulando la variable independiente, es decir, provocando deliberadamente una modificación de algún aspecto de la realidad (causa) para observar qué pasa con algún otro (variable dependiente o efecto) en la *investigación experimental* (o sea con diseños rigurosos o cuasiexperimentales) o sin hacerlo (*investigación no experimental*).

A su vez, los distintos estudios se relacionan con el tipo de preguntas que nos permiten responder, en la Tabla 1.4 se muestran algunas de ellas para los diseños cuantitativos. La última pregunta no es explicativa para algunos, pues sólo consideran explicaciones causales. Pero ya Hempel (1979) advirtió que las premisas de una explicación pueden contener muchos tipos de enunciados, además de los que establecen relaciones de

causa y efecto, aunque, por razones que no puedo tratar aquí, éstas han resultado las más atractivas para los investigadores.

Tabla 1.4. Estudios cuantitativos y tipos de preguntas más frecuentes que los orientan

Descriptivos	En el lapso de tiempo t y en la población N , ¿cuándo se dará el mayor porcentaje de Y ?, ¿cuántos sujetos presentan las características $Y_1, Y_2, \dots, Y_n$?, ¿cuál es la media de Y ?, ¿cuál es la mediana y la moda?, ¿resulta mayor la media de Y en el grupo x_1 que en x_2 ? Referido a un período determinado, ¿cómo variará Y ?
De contrastación de hipótesis (Explicativos)	¿ X influye en Y ?, ¿ X se relaciona con Y ?, ¿ W y Z influyen en Y ?, ¿Hay interacción entre W y Z en relación con Y ?, ¿se correlaciona X con Y ?

También debemos notar que suele usarse el término “correlacional” para mencionar a ciertos tipos de investigación descriptiva y a la que aquí considero no experimental (por ejemplo, Cohen y Manion, 1990). Pero este término por lo general crea confusiones, pues a veces la investigación correlacional se realiza con la finalidad de contrastar hipótesis, mientras que otras veces responde a una lógica descriptiva, por lo que considero que puede ser de uno u otro tipo.

Si ocurre lo primero, hablaré de estudios *no experimentales*, pues su nota distintiva es la ausencia de manipulación de variables; mientras que si ocurre lo segundo, hablaré directamente de *estudios descriptivos*.

La investigación no experimental tiene algunas limitaciones para analizar relaciones de causalidad, como ya veremos, pero esto no significa que carezcan de utilidad para ello, pues también tiene ventajas sobre los estudios cuasiexperimentales o rigurosos. Así, como regla general, los últimos representan la mejor posibilidad para contrastar este tipo de hipótesis, siguiéndole los cuasiexperimentales. Sin embargo, no podemos decir que un tipo de investigación es mejor que otro, sino simplemente que puede ser más o menos adecuado en función del objetivo que se propuso el investigador. Por ejemplo, si lo que quiere es poner a prueba un modelo que cuenta con muchas variables, en la mayoría de los casos será mejor una investigación no experimental multivariada que una cuasiexperimental o una experimental estricta, aunque en otro

momento estas últimas pueden resultar más convenientes para poner a prueba algunas hipótesis muy específicas del mismo.

En este trabajo me referiré a los diseños de investigación o tipos de estudio centrándome fundamentalmente en la validez interna. Pero antes finalizaré con la descripción sintética de las actividades del proceso de investigación no tratadas hasta ahora, para dar un panorama global del mismo.

La población y la muestra

Entiendo por *población* o *universo* al conjunto de todos los individuos de interés para nuestra investigación, que surge de un modo directo de los objetivos planteados. Cuando no podemos realizar el trabajo empírico con todos ellos, tomamos un subconjunto denominado *muestra*, que puede ser seleccionada básicamente de dos formas: la primera es aleatoria y la segunda no cuenta con este requisito.

En las muestras *aleatorias* todos los sujetos de la población tienen la misma probabilidad de ser seleccionados, dentro de ciertos márgenes de error permiten generalizar los resultados hallados en ellas al universo de interés e incluyen varios subtipos, según la técnica que se utiliza.

La primera se denomina *al azar simple* y para llevarla a cabo, en primer lugar, tomamos el listado de todos los individuos que componen la población y los numeramos; luego, con un procedimiento aleatorio (por ejemplo, introducimos en un bolillero tantas bolillas como individuos tiene el universo), seleccionamos el número de casos que hemos determinado.

La segunda forma se da cuando usamos el *muestreo sistemático*, que consiste en elegir un elemento cada k de la población, partiendo del primero seleccionado al azar. Por ejemplo, si ella cuenta con 1000 personas y decidimos trabajar con una muestra de 100, en primer lugar obtenemos al azar un número comprendido entre 1 y 10. Supongamos que fue el 7, éste es el primer individuo a escoger, luego le seguirán el 17, 27, 37, etc., hasta completar la cifra predeterminada.

El *muestreo estratificado* se utiliza cuando estamos en presencia de alguna variable de la cual sabemos que influye en el fenómeno de interés. Supongamos que sea la clase social²⁴. Formamos grupos (estratos) en función de esta propiedad (podrían ser más de una) y luego extraemos una muestra dentro de cada uno de éstos por un procedimiento al azar simple o sistemático. El muestreo estratificado puede ser *proporcional* o *no proporcional*.

“La muestra estratificada proporcional es aquella en la cual la fracción de muestreo es igual para cada estrato; si existen diferencias en las fracciones de muestreo, llamamos a ésta *no proporcional*” (Padua, 1987: 73).

Entre las formas de selección aleatorias, finalmente tenemos el *muestreo por conglomerados*. Por ejemplo, si nuestra población es una ciudad completa, podemos dividirla en manzanas (serían los conglomerados), elegimos al azar un número determinado de ellas y luego tomamos a todos los individuos de cada manzana o procedemos a extraer una cantidad menor con alguno de los métodos anteriores dentro de cada una.

Entre las formas *no probabilísticas* podemos identificar el *muestreo casual*, que consiste en seleccionar a los individuos que se hallan o transitan por un determinado lugar en un momento dado. Por ejemplo, nos ubicamos en una plaza y entrevistamos a un sujeto cada 10 de los que pasan por ese lugar. En el *muestreo intencional*, lo que hacemos es elegir a los individuos según la opinión de algún experto o de alguien que conoce muy bien a la población de interés, lo que permite estudiar casos típicos o con alguna característica de interés. En el *muestro por cuotas*, se extrae una proporción (cuota) correspondiente a la de la población en función de una o más variables determinadas. En Miles y Huberman (1994) pueden verse otras formas de muestreo no aleatorio que se utilizan en la investigación cualitativa, algunas de las cuales también suelen usarse en una aproximación cuantitativa.

Los instrumentos de recolección de datos

El instrumento de recolección de datos por lo general está compuesto por los materiales, las reglas de aplicación, las reglas de interpretación

24 En este caso, decimos que la clase social es la variable de estratificación.

de lo hallado y un sistema de registro. Los primeros, son aquellos que se le brinda o muestran al sujeto en el momento de recolectar el dato, por ejemplo, en el test de Bender se cuenta con las 10 tarjetas con las figuras que se deben copiar, papel y lápiz (a veces se suelen agregar otros materiales). La aplicación del instrumento de recolección de datos da como consecuencia la fijación de un hecho, tanto si el mismo es estructurado como si no lo es. Por ejemplo, véase el siguiente ítem modificado de Sposetti et al. (2004):

Durante el primer año de estudio, por lo general, cuando tenías que rendir un parcial estudiabas	
☐	Todos los días de la semana
☐	De cuatro a seis días por semana
☐	Dos o tres días por semana
☐	Un día por semana
☐	No estudiaba por lo general

Supongamos que el sujeto eligió la opción *Dos o tres días por semana*, en este caso, el hecho creado puede traducirse en el enunciado “en el tiempo t y el espacio e , el alumno a , afirmó que durante el primer año de estudio, *cuando no tenía que rendir parciales estudiaba dos o tres días por semana*”. Naturalmente, muchos detalles no se registran por obvios (como el tiempo y espacio). Incluso cuando se ingresan en la computadora, lo que se hace es codificarlos con lo cual todo el enunciado se puede reducir a un sólo carácter. Por ejemplo, si hemos decidido que la primera alternativa se codificará con un 1, la segunda con un 2, la tercera (que eligió el supuesto sujeto) con un 3 y así sucesivamente, en la base de datos ingresaremos este último número que equivale al enunciado antes mencionado.

Si el instrumento de recolección de datos es estructurado, a su vez, cada ítem tiene una escala. Según el modo más aceptado de clasificarlas estas pueden ser de cuatro tipos: nominales, ordinales, intervalares y de cocientes (o de razones). Un modo extremadamente simple de ver las tres primeras es el siguiente: “estamos ante una escala nominal si, para

dos unidades de análisis cualesquiera, sólo podemos decir que pertenecen a la misma o a diferente categoría. Una escala ordinal supone, además, poder establecer un orden entre ellas”, [... es decir que además de conocer si pertenecen a la misma categoría o a una diferente], podemos conocer cuál de las dos unidades de análisis tiene más cantidad de la propiedad que indaga el ítem²⁵. Si la escala es intervalar, podemos decir cuántas unidades, un caso tiene más que el otro de la propiedad o variable observada (Echevarría, 2008: 17)²⁶.

Para aquellos casos en que no contamos con opciones de respuestas predeterminadas el instrumento de recolección de datos también tiene por finalidad fijar ciertos hechos. Si por ejemplo, estamos haciendo un registro narrativo, todo lo que anotamos son los hechos registrados (podemos decir que los fijamos en un texto).

La aplicación a veces supone personal especializado (por ejemplo, observación etnográfica) aunque algunos instrumentos de recolección de datos sólo requieren de conocimientos generales, por ejemplo, tener el colegio secundario aprobado. Las instrucciones y las reglas de evaluación pueden ser más o menos explícitas. El sistema de registro puede ser una simple hoja y un lápiz, una lista de cotejo, un grabador, una cámara fotográfica, etc. Su función es fijar el dato, hacer por ejemplo, que un cierto segmento de conducta de un sujeto sea captado y perdure en el tiempo.

Los instrumentos de recolección de datos usados con más frecuencia en la investigación educativa son la *entrevista*, el *cuestionario* y la

25 Las propiedades de las escalas de los ítems, también valen para las variables compuestas por más de una dimensión, pues éstas también tienen alguna escala. Por otro lado, el término escala se lo usa con distintos sentidos, algunos de ellos son mencionados por Ander-Egg (2003): como sinónimo de nivel de medición (es el que uso aquí), para referirse a las escalas sociométricas y a las de medición de actitudes y opiniones. Las unidades de análisis en la investigación cuantitativa generalmente son individuos, aunque podrían ser grupos, instituciones, etc.

26 Las escalas de cocientes, de razón o proporcionales, cumplen con dos propiedades más aparte de las intervalares. “Si f_p y g_p son dos funciones numéricas para medir la propiedad P presente en los objetos de cierto dominio, o sea dos escalas diferentes para la misma magnitud P , decimos que son escalas proporcionales o de razón, si y sólo si para todo x , y se cumple: $f_p(x)/f_p(y) = g_p(x)/g_p(y)$ ” (Díez y Moulines, 1997: 118). Por esta razón, “sería más adecuado decir, no que una escala es una escala proporcional, sino que *un grupo* de escalas es un grupo de escalas proporcionales” (Ibídem, cursiva en el original). Hay que notar que estos autores usan un concepto muy restrictivo de escala, reservándolo para las intervalares y la proporcionales. La segunda propiedad que suele considerarse para las intervalares es la de tener “un punto cero real en su origen” (Siegel, 1991: 49).

observación. Con la primera, recabamos la información a través de una conversación con los sujetos investigados, que puede realizarse

“sobre la base de un formulario previamente preparado y estrictamente normalizado, a través de una lista de preguntas establecidas con anterioridad” (Ander-Egg, 2003: 90).

En este caso, decimos que la entrevista es *estructurada*. Si ocurre lo opuesto, es decir, no hemos establecido las preguntas y por lo tanto, tampoco las opciones de respuestas, la entrevista es *no estructurada*. Entre estos dos extremos, tenemos muchas opciones intermedias: podemos leer las consignas a todos los entrevistados del mismo modo, pero sin contar con alternativas de respuestas a priori y sin agregar otras preguntas, o incluir algunas en función de las respuestas de los sujetos o de las reflexiones que realizamos durante su implementación, etc.

El *cuestionario* se diferencia de la entrevista porque el investigador no interactúa con los sujetos que integran la muestra, sino que le da por escrito las preguntas que, a su vez, son contestadas de igual forma. Ander-Egg (2003) diferencia los casos en que el cuestionario se envía por correo y aquellos en los que se realiza a grupos (“de redacción colectiva”). Como no puede ser de otro modo, en él todos los ítems siempre se formulan de la misma manera, excepto que los investigadores de antemano tengan previsto cambiar algunos para determinados sujetos muestrales, suministrándoles formularios diferentes.

La *observación* consiste en registrar los acontecimientos relevantes, según van ocurriendo en un lapso y lugar determinados. También puede ser *estructurada* o *estandarizada* y, en este caso, debemos determinar de antemano qué aspectos observar, especificar sus dimensiones y seleccionar los instrumentos que se utilizarán (*ibidem*). Si se realiza un registro narrativo en forma secuencial de lo que va ocurriendo, decimos que es *no estructurada*, habiendo también opciones intermedias entre estos extremos.

Existen otras formas de clasificarla. Una de ellas, que es importante desde el punto de vista metodológico, se relaciona con el grado de participación que tiene el investigador. En uno de los extremos tenemos la *observación participante*, que se da cuando éste se incluye en la comunidad bajo estudio como un integrante más y, en el otro extremo, se halla la *observación no participante*, en la que realiza su trabajo pero sin intervenir en las actividades que los actores sociales están desarrollan-

do. Algunos incluso diferencian entre observador participante y participante observador. En el primer caso, el investigador elige a uno o más integrantes del grupo de personas que le interesa estudiar y le pide que realice el registro de la información; en el segundo, se inserta en este grupo con la intención deliberada de concretarlo.

Existen otros instrumentos de recolección de datos, aunque es difícil mencionar todos los tipos, pues no existe un único criterio para clasificarlos. Incluso algunos los llaman de un modo diferente o los consideran subtipos de otro.

A veces puede hacerse la investigación con *datos secundarios*, que son aquellos recolectados para otra finalidad o investigación (diferente al análisis de contenido en que el investigador selecciona los materiales a usar). Ejemplos de ellos pueden ser: Registro de alumnos de universidades, datos de colegios, hospitales, consultorio profesional, Censos. Su inconveniente principal es que no siempre los datos se adecuan a los objetivos, y generalmente estos se deben subordinar a lo que es posible lograr con ellos. Su principal ventaja no es una cuestión menor: tenemos los instrumentos de recolección de datos contruidos y el trabajo de campo realizado. La investigación con datos secundarios suele tener problemas de validez y confiabilidad (no siempre sabemos cómo se recolectaron, falta conocimiento de lo que significan las categorías, las preguntas que se hicieron, etc.). También suelen surgir problemas con las transformaciones de archivos (no siempre es factible grabarlos en formatos que se puedan importar a nuestros programas). Por todo ello, se recomienda chequear la calidad de los datos antes de iniciar un proyecto de esta naturaleza.

El procesamiento de los datos

Al presentar el proyecto, sólo debemos explicitar los procedimientos de reducción de los datos que pensamos utilizar. Luego de recolectados, los mismos serán aplicados, aunque antes debemos depurarlos e ingresarlos en la computadora.

A veces se suele considerar una misma actividad al procesamiento y análisis de los datos. Aquí, siguiendo a Sabino (1996), las considero dos cosas diferentes: por un lado el investigador procesa los datos y, por otro lado, los analiza, pues la información procesada, si bien

“tiene un valor inestimable (...), no obstante, (...) no nos ‘habla’ por sí misma, no es capaz por sí sola de darnos las respuestas deseadas hasta tanto no se efectúe sobre ella un trabajo de análisis e interpretación” (: 203).

Con los estudios descriptivos, apelamos generalmente a tres tipos de estadísticos: las *medidas de tendencia central* (nos dan una idea del valor de alguna variable para la mayoría de un grupo de sujetos), las *medidas de dispersión* (nos indican cuán dispersos o cuán próximos están los valores registrados para cada sujeto respecto de una medida de tendencia central) y las *tablas de frecuencias* (a veces suelen presentarse en forma de gráficos y nos informan sobre la cantidad de casos que incluye cada categoría o valor de una variable). Dentro de las medidas de tendencia central, las más usadas son la media aritmética, la mediana y la moda; y entre las medidas de dispersión, podemos mencionar a la desviación estándar.

Estas medidas pueden usarse para describir una variable sin relacionarla con otra; pero, si esto es lo que queremos hacer, con ellas podemos desarrollar la prueba de significación, que por lo general se utiliza en estudios explicativos, aunque también se puede implementar en los descriptivos. Ésta se complementa con el coeficiente de correlación, que nos

“da información sobre la *intensidad* de la relación [entre variables], mientras que la prueba de significación lo hace sobre la *seguridad* de la relación. La prueba de significación indica con qué seguridad puede afirmar el investigador la existencia de relación. El coeficiente indica en qué medida guardan una relación mutua las dimensiones (sin dar información sobre la seguridad con que cabe afirmar tal relación)” (Mayntz, et. al. 1975: 154).

Con el *cruzamiento*, analizamos en qué medida las variables de un cuadro de dos o más entradas están asociadas, por lo que es una técnica más imprecisa que el coeficiente de correlación. Si se considera más de una variable independiente o dependiente, tendremos que apelar a los análisis multivariados.

Por la sencilla razón de que aún no han sido recolectados, en el proyecto no podemos incluir el análisis, interpretación y discusión de los datos, lo que comienza luego del procesamiento y consiste en darles un sentido desde los objetivos que nos propusimos y las preguntas que

formulamos, resaltando lo más importante de ellos en función de las últimas. Si, por ejemplo, nuestra pregunta fue “¿De dónde provienen los ingresantes de la universidad *X* en el año 2015?”, resaltamos cuáles son las mayores frecuencias en función de esta variable, si provienen mayoritariamente de una provincia determinada, de la misma localidad en que está la universidad, de otras vecinas, etc. Si el objetivo es explicativo, fundamentalmente debemos ver si la hipótesis planteada se corroboró o se refutó y, en todos los casos, es importante realizar una valoración de lo hallado, comparándolo con otros estudios, formulando nuevas preguntas o recomendaciones para intervenir sobre la realidad. Estos últimos aspectos suelen incluirse dentro de lo que en los informes se denomina “discusión”.

El cronograma de actividades

El cronograma de actividades, como su nombre lo indica, tiene por finalidad especificar para cada mes qué actividades se cumplirán (puede hacerse semanalmente o tomando otra unidad de medida si se lo cree necesario). Es de gran ayuda para los investigadores, pues los guía y les permite observar el grado de avance del proyecto, y al mismo tiempo, facilita a los evaluadores conocer su factibilidad y coherencia interna (el grado en que las distintas piezas encajan entre sí). El cronograma puede concretarse a través de un cuadro de doble entrada en el que ubicamos las actividades en las filas y los meses en las columnas. El Cuadro 1.3 muestra un ejemplo hipotético para un estudio descriptivo que se desarrolla en un año. Puede verse que algunas actividades se superponen, es decir se llevan a cabo en forma paralela, como es la recolección y procesamiento de los datos. Supongamos que se trata de un estudio longitudinal en el que el trabajo de campo se desarrollará entre los meses 4 al 6. En este caso, ni bien se cuenta con la información del primero o parte de ella, ya se puede estar ingresando en la computadora, con lo que se va ganando tiempo, y lo mismo ocurre con otras actividades.

Cuadro 1.3. Cronograma de actividades por mes para un proyecto descriptivo hipotético de carácter anual

Actividad	Mes									
	2	3	4	5	6	7	8	9	10	11
Prueba de los instrumentos de recolección de datos	X									
Corrección de los instrumentos de recolección de datos	X	X								
Recolección de los datos			X	X	X					
Procesamiento de los datos				X	X	X	X	X	X	X
Redacción del informe final							X	X	X	X
Presentaciones en eventos científicos							X	X	X	X

Los recursos disponibles y solicitados

En este ítem se deben detallar con la mayor precisión posible aquellos recursos materiales y humanos de los que se dispone y que se requieren, pues esto permitirá a los evaluadores conocer si se justifican los fondos solicitados para la investigación. Si se trata de un proyecto en el que no se solicita subsidio (como suele ocurrir, por ejemplo, para un trabajo final de grado), obviamente no se incluirá el último ítem, pero se debe mostrar, a través del primero, que el mismo es factible con los elementos disponibles.

Hasta aquí, he tratado de un modo muy sintético todos los aspectos que se relacionan con el proceso y el proyecto de investigación; a partir del capítulo siguiente me centraré en el que constituye el tema central del presente libro: los diseños de investigación.

Ejercicios

1. Haz una primera lectura del primer capítulo y buscá en alguna base de datos (por ejemplo, www.redalyc.org), en bibliotecas, etc. e identificá un proyecto y un informe de investigación (seleccioná, en la medida de lo posible, dos publicaciones breves). Como variación, el docente puede brindarle a los alumnos ambas publicaciones.
2. Compara los escritos, ¿qué apartados incluye cada uno?, ¿todos consideran los mismos apartados?, si la respuesta es negativa, ¿qué diferencias hay entre ellos?
3. ¿Qué forma de citar usan según los estilos que presenta Rusconi (2003)²⁷?
4. Señala en el margen cada uno los elementos que incluyen y se describen en el Capítulo 1 (antecedentes empíricos, problema, supuestos, definiciones, hipótesis, etc.).
5. En ellos, ¿qué conocimientos dan los autores por “satisfactoriamente corroborados” (lo presentan como no problemático o aceptado por la comunidad científica). ¿Citan estudios para justificarlos? Subraya las citas. ¿Considerás que las citas están bien referenciadas (es inequívoco el texto de donde se extrajo, se aplican las normas para citar correctamente)?
6. Si las autoras o autores explicitan alguna o algunas dudas respecto al área de conocimiento que trata el informe o proyecto, transcribí en forma completa un párrafo en que lo hacen (incluí la página).
7. Si los objetivos están explícitamente enunciados, transcribilos. Si no lo están, enuncialos con sus propias palabras.
8. Si están explicitados en el informe, transcribí al menos dos supuestos del marco teórico. ¿Citan estudios o investigaciones en que los fundamentan? Si tu respuesta es afirmativa, transcribí las citas.

²⁷ Si no se cuenta con este material se puede utilizar alguna otra publicación donde se tratan las convenciones de la escritura científica

9. Si el o los autores las incluyen, transcribí dos definiciones (explicitá las páginas en que lo hacen).
10. Estas definiciones, ¿son operacionales?, ¿por qué?
11. ¿En algunas de las respuestas a las preguntas anteriores, incluiste el planteo del problema de investigación en tus respuestas?, si tu respuesta es afirmativa, ¿en cuál?
12. ¿Cuál es el o los problemas prácticos o sociales a los que se vincula cada proyecto o informe de investigación?

Siguiendo las convenciones mencionadas por Rusconi (2003), debés realizar (agregar la bibliografía o referencias de todas las citas):

13. Una cita directa de un *libro* con el sistema Autor/año de edición
14. Una cita indirecta con el sistema Autor/año de edición
15. Una cita directa de:
 - 15.1. Un capítulo de libro,
 - 15.2. Una revista académica,
 - 15.3. Una presentación en congreso,
 - 15.4. Un documento no publicado,
 - 15.5. Una revista comercial o de divulgación,
 - 15.6. Un documento electrónico.

CAPÍTULO 2²⁸

DISEÑOS RIGUROSOS

La validez interna

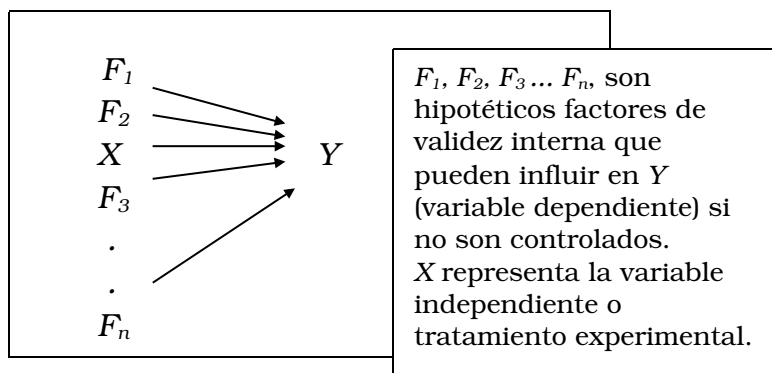
Desde el punto de vista metodológico, dos cuestiones resultan importantes a la hora de seleccionar el diseño de investigación. Una de ellas es la validez interna, la otra es la validez externa. Antes de presentar las distintas alternativas, veremos brevemente la primera, aunque por la importancia que tiene este concepto, lo retomaré en varias oportunidades a lo largo del trabajo. Sobre la validez externa, sólo haré unas pocas observaciones, pues, debido a la naturaleza introductoria de este libro, el abordaje que presento se centrará fundamentalmente en la primera.

Podemos decir que un *experimento* es un suceso deliberadamente creado para conocer cuál es el efecto que produce una variable sobre

28 Este capítulo y los dos siguientes están basados fundamentalmente en Shadish, et al. (2002), por lo que pueden usarse como una introducción a ellos, aunque también se toman ideas de Campbell y Stanley (2005) y otros autores. No obstante, el análisis de la validez en estos dos trabajos es extremadamente complejo en comparación con lo que se presenta aquí, que sólo pretende dar una idea introductoria a este tema. Hay que notar que no utilizo exactamente la misma terminología, pues debí adecuarla para integrar los diseños descriptivos que ellos no consideran. Por ejemplo, a los estudios que aquí llamo *rigurosos* ellos los mencionan como *experimentales*.

otra. A la primera, llamada *variable independiente* (causa), se le hace asumir un valor distinto al que tiene habitualmente (o simplemente se la hace variar entre al menos dos de sus categorías o valores posibles) con la finalidad de observar si se produce algún tipo de modificación en otra variable, llamada *dependiente* (efecto)²⁹. Por ejemplo, el investigador quiere saber si un nuevo método de enseñanza de la lectoescritura (que llamaré *M1* por razones de comodidad) es mejor que el que se usa actualmente (*M2*) en un determinado colegio. En este caso, la variable independiente es el método de enseñanza y los valores que puede asumir son *M1* y *M2*. Podemos enseñarle al primer grado *A* del turno mañana con *M1*, mientras que al primer grado *B* del turno tarde con *M2* y luego de transcurridos 6 meses tomarle una prueba de lectoescritura a ambos grupos. Supongamos que el grupo que aprendió con *M1* registra una media superior al que lo hizo con *M2*. ¿Significa esto que el primer método es mejor que el segundo? No podemos estar seguros, ya que es posible que ambos grupos no hayan sido equivalentes antes de la enseñanza, o aún siéndolo puede haber otras condiciones, como el horario por ejemplo, que provocó la diferencia en el aprendizaje. A estas condiciones, ajenas a la variable independiente y que pueden influir sobre la dependiente, las llamamos *factores de validez interna* (ver Figura 2.1).

Figura 2.1. Factores de validez interna



²⁹ En realidad pueden ser más de una variable independiente, como así también más de una dependiente. Por razones didácticas hago la exposición suponiendo se cuenta con una sola de cada clase.

¿Cómo controlar, entonces, estos factores ajenos al tratamiento experimental (en este caso el método de enseñanza) y que por sí mismos pueden provocar una diferencia entre los grupos? Existen distintos modos de control: el estadístico, el apareamiento, la realización de tratamientos múltiples, el agregado de distintas mediciones en diseños cuasiexperimentales y la aleatorización o distribución al azar

En el ejemplo hipotético presentado, un modo de implementar la última sería formando dos subgrupos con el grado de la mañana únicamente, asignando aleatoriamente a cada sujeto al grupo control o experimental. Estrictamente hablando, no serían totalmente equivalentes, ya que siempre persistirá una diferencia entre ellos (que será menor cuanto mayor sea la muestra seleccionada, aunque no estamos ante una función lineal). Pero por suerte, la estadística nos brinda los medios para saber si las diferencias observadas en el aprendizaje se deben al azar o podemos atribuírsela a la variable independiente (el método de enseñanza). Esta forma de control por aleatorización es la más elaborada que conocemos y los diseños que cuentan con ella son denominados aquí *experimentales estrictos o rigurosos (usando ambos términos como sinónimos)*.

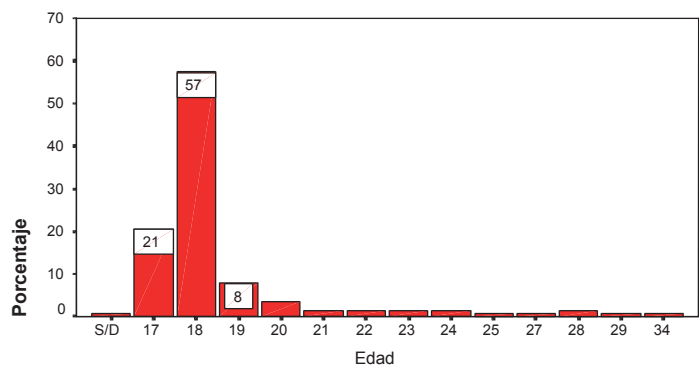
Los estudios con control riguroso presentan la enorme dificultad de requerir la distribución al azar, cosa que no siempre es posible, además de provocar una artificialidad en el experimento que a veces afecta la *validez externa* (es decir, a la posibilidad de generalizar al universo de interés a partir de la muestra utilizada).

Dada la importancia que tiene comprender acabadamente lo que significa la distribución al azar y el modo en que se vincula con la validez interna, veamos con un ejemplo este aspecto con más detalle. Para tal fin, se tomó un grupo de alumnos ingresantes a la Facultad de Ciencias Económicas de la U.N.R.C. del año 1999 y se lo dividió al azar en dos subgrupos. Después se obtuvieron los gráficos 2.1 al 2.5, que muestran las frecuencias según edad, sexo y número de hijos de ambos. Puede apreciarse que, si bien no se obtuvieron idénticos valores en las tres variables, éstos son similares. Los dos grupos muestran la mayor cantidad de alumnos con 18 años, luego le sigue 17, en tercer lugar tenemos la edad de 19 años y luego 20, con bajas frecuencias en las restantes (ver Gráficos 2.1 y 2.2). Con el sexo y el número de hijos, las distribuciones también se parecen (ver Gráficos 2.3, 2.4 y 2.5). La distribución al azar,

si bien no logró equiparar en forma total a ambos grupos, en términos generales los igualó en las tres variables que analizamos.

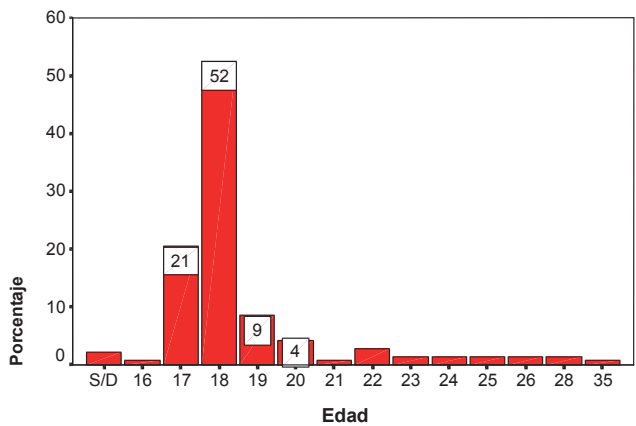
Es posible lograr la aleatorización en los tres aspectos siguientes: 1) selección de la muestra, 2) distribución de los casos (que generalmente son sujetos) en los grupos, y 3) determinación de cuáles serán los grupos de control y experimentales (Kerlinger, 1990). No obstante, cualquiera de ellos podría no realizarse al azar, aún cuando esto suceda con los otros dos. En este trabajo consideraré *riguroso* a todo estudio en que se da al menos el punto 2) y que a su vez, supone hacerlo también con el 3.

Gráfico 2.1. Cantidad de alumnos por edad en el grupo 1



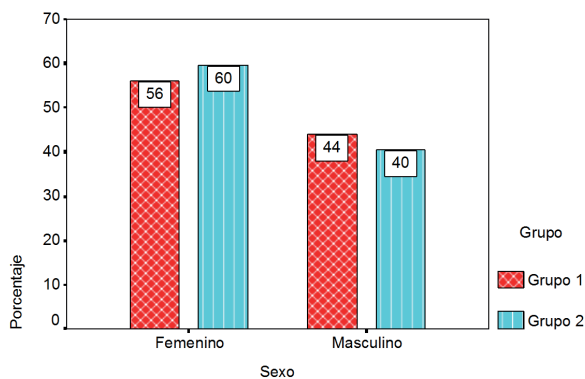
Fuente: Centro de Cómputos, Facultad de Ciencias Económicas

Gráfico 2.2. Cantidad de alumnos por edad en el grupo



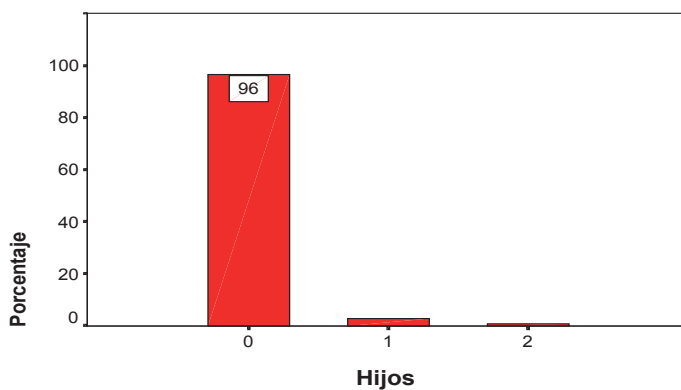
Fuente: Centro de Cómputos, Facultad de Ciencias Económicas

Gráfico 2.3. Cantidad de alumnos en cada grupo, según sexo (en porcentajes)



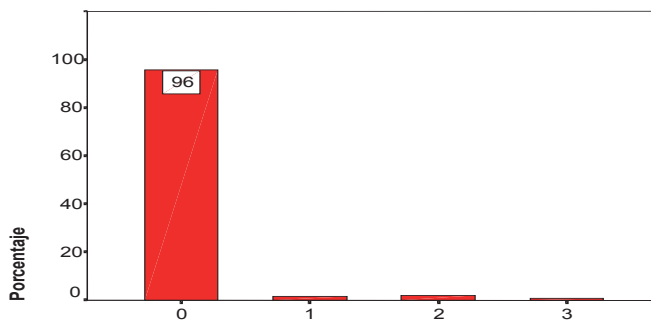
Fuente: Centro de Cómputos, Facultad de Ciencias Económicas

Gráfico 2.4. Cantidad de alumnos según número de hijos en el grupo 1



Fuente: Centro de Cómputos, Facultad de Ciencias Económicas

Gráfico 2.5. Cantidad de alumnos según número de hijos en el grupo 2



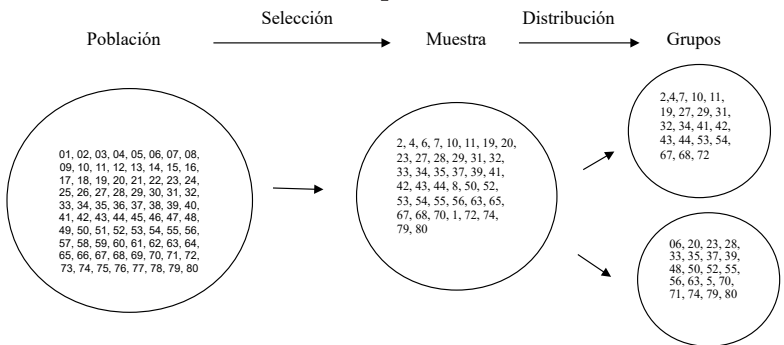
Fuente: Centro de Cómputos, Facultad de Ciencias Económicas

De todas estas consideraciones, surge una pregunta central para el tema que nos ocupa: ¿cuántas son las variables a controlar para lograr una cierta validez interna? Éstas son prácticamente infinitas, pero la aleatorización (asignación al azar de los sujetos a cada grupo) permite igualar los grupos en todas ellas (dentro de ciertos límites como dije más arriba), por eso, como norma general los diseños rigurosos tienen mayores posibilidades de control de la validez interna. Más adelante matizaré esta afirmación, pues deben tenerse presente las condiciones contextuales en que se lleva a cabo el estudio. Hablando estrictamente, en realidad debíamos decir que la aleatorización permite que los grupos sean más similares a medida que se aumenta el tamaño de la muestra, lo que, a su vez, es explicado por la teoría de las probabilidades aunque su tratamiento escapa ampliamente a los objetivos del presente trabajo.

La aleatorización en los diseños rigurosos

La Figura 2.2 muestra un caso hipotético en el que, partiendo de una población de 80 sujetos, se *seleccionaron* al azar 40 y luego se formaron (se *distribuyeron* en) dos grupos de 20 (por igual procedimiento). Si, además, se determina también de esta forma cuál será el grupo experimental y cuál el de control, cumpliremos con los tres requisitos. El segundo, la distribución, es el más importante de todos, pues es el que se vincula de un modo directo a la validez interna, sin la cual de nada nos serviría la validez externa.

Figura 2.2. Proceso de selección de la muestra y de formación de grupos para estudios experimentales.



En la Figura puede verse que mediante la *selección* obtenemos una muestra a partir de la población y, con el proceso de *distribución*, formamos grupos con la muestra. Tanto uno como otro proceso podrían no ser aleatorios, aunque si se logra al menos el segundo, considero que el diseño es *riguroso*.

Es muy común que se confundan la selección (de la muestra) con la asignación (de casos a grupos). Por la primera, obtenemos una muestra, esto es, la totalidad de casos con los que haremos la investigación, y con la segunda, a partir de la muestra obtenemos grupos. Incluso, a veces estos pasos se cumplen en forma abreviada, lo que puede aumentar la confusión. Por ejemplo, supongamos que un investigador piensa hacer un estudio en una escuela primaria que cuenta con 8 divisiones en primer grado. Decide seleccionar dos para el mismo y lo hace al azar simple.

Además, de antemano establece que el primero que extrae sea el grupo de control y el segundo el experimental. Supongamos también que introduce 8 bolillas en un bolillero y luego extrae dos al azar. En esta situación, cumple los tres pasos en uno solo y pareciera que los tres aspectos se realizan aleatoriamente. Sin embargo, no es así, pues en realidad está tomando dos *grupos naturales*, o sea, formados con otra finalidad, y se incluyen en forma completa (tal como están en la situación natural) a tratamiento o control, por lo que la asignación de los sujetos a los grupos no se hizo al aleatoriamente. *Para que esto suceda, el investigador debe romper los grupos naturales y formarlos de modo tal que todos los casos tengan la misma probabilidad de ir a uno u otro grupo.*

Podemos verlo de otro modo que tal vez clarifique el concepto: supongamos por comodidad que cada una de las divisiones cuenta con 40 alumnos; si extraemos dos divisiones, la cantidad de objetos elegidos al azar es dos divisiones y no 80 alumnos. Si bien las divisiones se eligieron al azar, no podemos aplicar la estadística como si fueran dos grupos de 40 casos y como sabemos, no sería aplicable una prueba estadística con un sujeto por cada grupo, pues en realidad los elegidos al azar son dos grupos y no los alumnos.

Algunos incluso sostienen que, además de cierta forma del diseño, se debe cumplir con el requisito del doble ciego. Esto significa que ni los sujetos involucrados ni el experimentador conocen quiénes están sometidos al tratamiento y quiénes al placebo (control). Esto último es

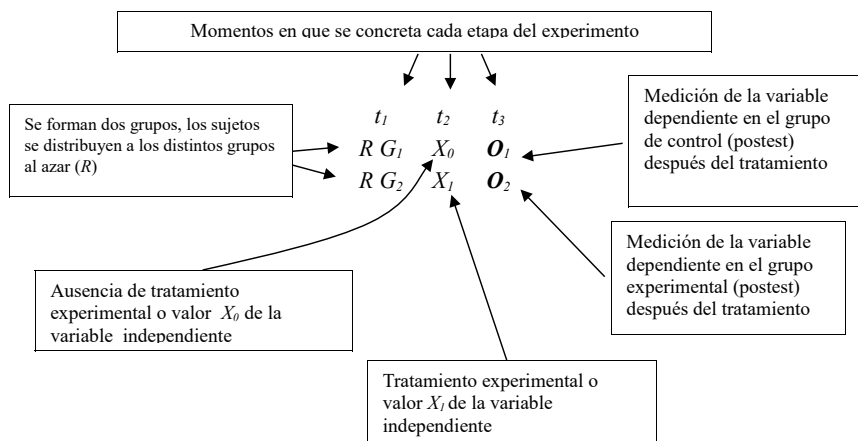
imposible de lograr en muchas situaciones en Ciencias Sociales, lo que representaría una limitación insuperable para aplicar este tipo de metodología. Volveré sobre este punto más adelante, en donde relativizaré esta afirmación.

Por ahora presentaré algunos de los diseños posibles, seleccionando aquellos que mejor permitirán explicar los conceptos involucrados. El lector deberá tener presente que siempre podrá modificar alguno de ellos mediante la substitución de alguna de sus partes (lo que a veces se hace para ganar en simplicidad), el agregado de otras o a travéz ambas cosas.

Diseño de dos grupos, con distribución al azar y postest solamente

En el ejemplo mencionado más arriba, referido a la lectoescritura, si en vez de tomar los grupos naturales (es decir, tal como están constituidos), el investigador decide conformarlos distribuyendo a los sujetos en cada uno de ellos por algún procedimiento de aleatorización, tendríamos el diseño de dos grupos con distribución al azar y sólo postest, que presento en el Diagrama 2.1. Recordemos la simbología presentada en el Capítulo 1. t_1 , t_2 y t_3 indican los momentos (o tiempos) en que ocurre cada etapa. En la primera (t_1), se conformaron dos grupos (G_1 , el grupo de control y G_2 , el grupo experimental) *al azar*, lo que se representa con la R (recuérdese también que es la inicial de la palabra inglesa “random”, que en castellano podemos traducir como “azar”). *En cada columna, incluimos lo que coincide en el tiempo, y en filas hacemos lo propio en relación a los grupos (por ello, en esta simbología se usa una fila por cada grupo)*. Cada “ O ” representa alguna medición u observación de la variable dependiente realizada a todos los sujetos del grupo, aunque luego se usa algún estadístico que lo representa (es común usar la media, pero pueden ser otras medidas). En este diseño no es necesario usar dos dígitos, pues al darse las observaciones en un solo momento, el primero permanece constante (sin él diferenciamos perfectamente las observaciones a realizar).

Diagrama 2.1. Diseño de dos grupos aleatorios y sólo posttest



Las observaciones se efectúan sólo luego del tratamiento, por lo que hablamos de *posttest*. X_1 denota la existencia de algún tratamiento experimental y X_0 la ausencia de él o lo que se cumple habitualmente (en el ejemplo de la enseñanza de la lectoescritura, representan el método nuevo y el habitual, respectivamente). En una primera aproximación, podemos interpretar que la diferencia entre O_1 y O_2 es provocada por la diferencia en los tratamientos (X , la *variable independiente*), porque la distribución al azar nos garantiza que los grupos resulten equivalentes antes de la realización del experimento (*más adelante veremos que esta afirmación debe tomarse con mucha cautela*).

Diseño de dos grupos, con distribución al azar, pretest y posttest

Es similar al anterior (ver Diagrama 2.2), pero agrega una medición de la variable dependiente antes del tratamiento experimental, que se denomina *pretest*. Permite más comparaciones, pues además de hacerlo con los posttest (en este caso O_{21} y O_{22}), podemos calcular $O_{21} - O_{11} = Dc$, $O_{22} - O_{12} = De$, donde De y Dc representan la *ganancia* (es decir la *variación o diferencia entre el pretest y el posttest*) en los grupos experimental y de control, respectivamente. Finalmente, se puede conocer la diferencia entre De y Dc , que indica la medida en que la variable independiente influye

sobre la dependiente. También es importante recordar que en el subíndice el primer dígito representa la observación (en este caso, 1 antes del tratamiento y 2 luego de él) y el segundo, el grupo.

Este diseño se puede ampliar en varios sentidos. En primer lugar, agregando más grupos si la variable independiente tuviera más de dos valores (ver Diagrama 2.3 para cuatro valores de X).

En segundo lugar, agregándoles un segundo posttest³⁰. Esto es particularmente importante cuando, además de la intensidad del efecto de la variable independiente, se desea conocer la persistencia del mismo en el tiempo. En este caso, también suele decirse que se le agrega un posttest diferido. El tiempo que transcurre entre uno y otro posttest dependerá fundamentalmente del modelo que el investigador está contrastando (ver Diagrama 2.4).

Diagrama 2.2. Diseño de dos grupos aleatorios con pretest y posttest

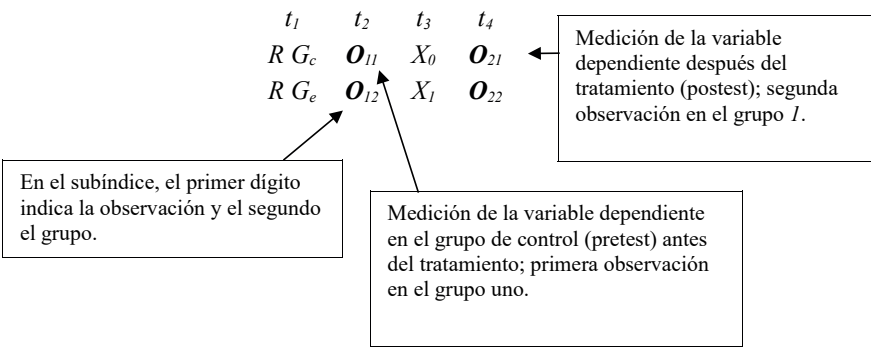
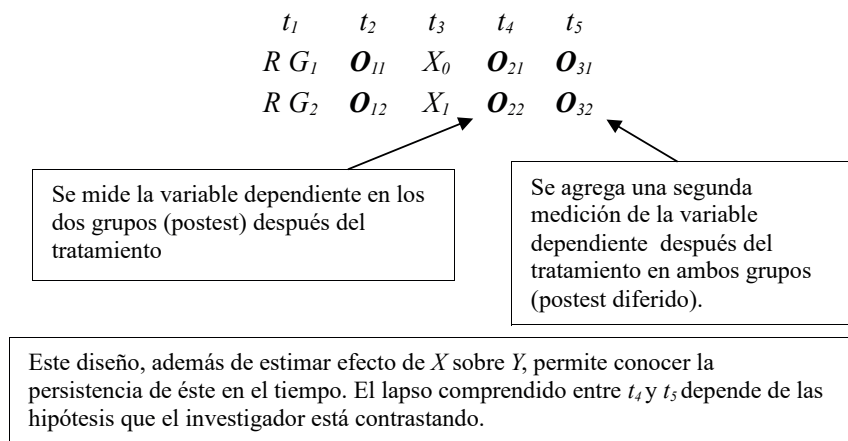


Diagrama 2.3. Diseño de cuatro grupos aleatorios con pretest y posttest

t_1	t_2	t_3	t_4
$R G_1$	O_{11}	X_0	O_{21}
$R G_2$	O_{12}	X_1	O_{22}
$R G_3$	O_{13}	X_2	O_{23}
$R G_4$	O_{14}	X_3	O_{24}

³⁰ También puede realizarse con tres, cuatro o más grupos.

Diagrama 2.4. Diseño riguroso de dos grupos con pretest y doble posttest



Como ejemplo, consideremos las siguientes hipótesis:

H1. Un programa de modificación de la conducta produce aprendizajes más rápidos que uno constructivista.

H2. Se da una menor persistencia en el tiempo con un programa de modificación de la conducta, respecto del recuerdo de los contenidos asimilados.

Así, el primer posttest permitirá dar cuenta de *H1*, pero para saber qué ocurre con *H2*, necesariamente debemos incluir un segundo posttest. Incluso, podría ser necesario agregar más de uno, por ejemplo, si modificamos *H2* como sigue:

H2. La diferencia en los recuerdos de los contenidos asimilados con un modelo de enseñanza constructivista y uno de modificación de la conducta, será mayor cuando mayor sea el tiempo en que finalizaron las actividades de aprendizaje.

Diseño de cuatro grupos de Solomon

Entre los diseños de control riguroso, el de cuatro grupos de Solomon goza de mucho prestigio. Representa una combinación de los dos que

vimos en primer lugar, pues se realiza el experimento con pretest y sin él. Es particularmente recomendado cuando quedan dudas del control del efecto del pretest o la interacción que existe entre éste y algún otro factor de validez (ver Diagrama 2.5).

Para muchos autores, los diseños de control total representan la máxima perfección posible. Dejando de lado esta cuestión, por cierto discutible, pues no en todos los casos son los más recomendables, a veces si siquiera se pueden aplicar, ya que suponen la enorme dificultad de la distribución al azar, algo que no siempre está a nuestro alcance³¹. Ante esta alternativa, debemos apelar a los estudios cuasiexperimentales o no experimentales, en el próximo capítulo, continúo con los primeros.

Diagrama 2.5 Diseño de cuatro grupos de Solomon

t_1	t_2	t_3	t_4
$R G_1$	O_{11}	X_0	O_{21}
$R G_2$	O_{12}	X_1	O_{22}
$R G_3$		X_0	O_{23}
$R G_4$		X_1	O_{24}

Ejercicios

1. De los trabajos presentados en el Apéndice B, ¿cuál o cuáles son rigurosos?
2. Si alguno es riguroso, realizá el diagrama del diseño.

³¹ Además, la artificialidad que se da con ellos puede interferir en las posibilidades de generalizar, lo que representa un objetivo básico en esta línea de investigación. Aquí tenemos implicados importantes problemas epistemológicos y metodológicos, aunque por razones de espacio, no serán tratados en este trabajo.

Capítulo 3

Diseños cuasiexperimentales

Diseños cuasiexperimentales sin grupo de control

Los estudios cuasiexperimentales, por las posibilidades de aplicación que implican, son muy importantes para las Ciencias Sociales, aunque difíciles de caracterizar metodológicamente. Son más fáciles de implementar porque consideran *grupos naturales*, es decir, los toman tal como están constituidos y se asignan a los distintos tratamientos de ese modo (por ejemplo, el primer grado *A* y el primer grado *B* en un colegio, o el turno de la mañana y el de la tarde). Además, buscan el control de los factores de validez interna, no por su neutralización como los rigurosos, sino por medio de su medición o a través de la compensación, aspectos que veremos al tratar cada uno de ellos.

Por otro lado, aunque no me propuse tratar el tema de la validez externa, puesto que en general los diseños rigurosos producirán más interferencias en la realidad investigada que los que no lo son, es menos probable que los cuasiexperimentales se vean afectados por ella.

Los diseños que veremos en primer lugar se caracterizan por contar con un solo grupo (o un solo sujeto), por lo que la comparación se hará

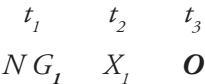
entre lo que ocurre antes y después del tratamiento, o directamente se compara con un grupo de referencia implícito, como en el primero que presento. Se trata del *diseño de un grupo y sólo posttest*.

Si bien comienzo con el que podría presentar el mayor número de amenazas de validez interna para avanzar gradualmente hacia aquellos en que éstas son menos probables, no debe suponerse un orden estricto en cuanto a este punto. Por ejemplo, el *diseño de un grupo, posttest y doble pretest* se expone antes que el *diseño con pretest y posttest* con variable dependiente no equivalente, pero eso no significa que necesariamente el segundo cuente con menos amenazas que el primero. En todo caso el lector debe recordar una idea central acerca de la consideración de la validez que aún a riesgo de ser reiterativo mencionaré varias veces a lo largo del presente trabajo: *la validez no puede plantearse exclusivamente a partir de la forma del diseño, sino que también deben considerarse las condiciones contextuales y en general todos los aspectos que involucra cada investigación específica*.

Diseño de un grupo con posttest únicamente

El diseño experimental más elemental es el de *un grupo y solo posttest* (Diagrama 3.1). En él, se toma un grupo, se lo somete a un tratamiento y luego se realiza la medición de la variable dependiente. Por ejemplo, se propone un curso de estrategias de aprendizaje con el cual suponemos que los alumnos universitarios tendrían menos dificultades para comprender textos científicos. Para comprobarlo, un docente toma un grupo de alumnos determinado y les propone que asistan a dicho curso. Luego de que los estudiantes lo han hecho, les toma una prueba de comprensión lectora y comprueba que son mayoría quienes la superan satisfactoriamente. ¿Podemos estar seguros de que manifestaron tener buena comprensión de los textos científicos por el curso que tomaron?

Diagrama 3.1. Diseño de un grupo y sólo posttest



La respuesta es no, pues entre muchos de los problemas que tiene, al carecer de pretest, no se puede comparar la variable dependiente antes y después del tratamiento. Es decir, no se puede constatar si efectivamente hubo una modificación en los alumnos que tomaron el curso (podría suceder que ya tuvieran buena comprensión lectora antes de hacerlo).

Shadish et al. (2002) nos hablan de las *amenazas de validez*, entendidas como “razones por las que una inferencia puede ser incorrecta” (: 512), inferencias que en los estudios de contrastación de hipótesis causales nos llevaría a creer en una relación causal que en realidad no existe. En esta definición se pone de manifiesto que ellos consideran que no necesariamente una inferencia es incorrecta porque el diseño tenga una forma determinada.

Las *amenazas de validez* también pueden verse como explicaciones alternativas (a la variable independiente) de los resultados hallados. Esto es, como otras causas que pueden influir en la variable dependiente en el mismo sentido. En el ejemplo recién mencionado, la *historia*, esto es, los sucesos que se desarrollan en forma paralela al experimento son una explicación alternativa posible. Por ejemplo, puede suceder que los alumnos hayan recibido explicaciones de ese tema en otra materia y por ello manifiesten luego un buen nivel en la prueba de comprensión.

Shadish et al. (2002) sostienen que “la lista de las amenazas de validez interna relevantes es enorme [si es que pudiera hacerse]. Sin embargo, la verosimilitud de una amenaza es siempre contextual, depende de las características que presenta el diseño, del conocimiento extra estudio de las amenazas, y del patrón de los resultados observados del estudio. Por lo tanto, las *posibles* amenazas a la validez no siempre son *plausibles*” (Shadish et al. 2002: 139, cursivas en el original).

“Esto significa que si se usó un diseño que deja serias dudas acerca de si alguna explicación alternativa es posible, el conocimiento que tiene el investigador de esa amenaza para ese caso particular, sumado a los resultados observados y al conocimiento acerca de cómo se realizó el experimento, puede permitirle que la descarte y recíprocamente, si fuera un diseño ‘fuerte’ en validez, estos elementos podrían hacernos dudar de que se haya observado realmente la relación causal hipotetizada” (Echevarría, 2011b, 184).

En el ejemplo recién mencionado, podría conocer cuáles son las materias que los alumnos están cursando junto a la suya, o tomarle un

cuestionario luego de que hayan contestado a la prueba de comprensión indagando posibles variables que explicarían los resultados observados.

De todos modos, en general el diseño de un grupo y solo posttest es poco recomendable por las amenazas de validez que tiene. Shadish et al. (2002) expresan muy claramente cuáles son sus alcances:

Puede ser útil cuando se conoce mucho sobre el comportamiento de la variable dependiente en determinadas situaciones en que se observa un efecto que es lo suficientemente intenso como para descartar ciertas amenazas (“el diseño tiene mérito en el caso poco frecuente, en el que muchos conocimientos específicos de fondo existen acerca de cómo se comporta la variable dependiente” (Shadish et al. 2002: 107). Estos autores ofrecen el siguiente ejemplo: sabemos que los niños prefieren mirar televisión o jugar con la computadora antes que leer libros, y estamos ante un grupo típico en ese sentido. Un docente presenta una propuesta para favorecer el interés de los pequeños por la lectura. Si se observa un efecto notable, podemos interpretar el resultado del posttest como producido por el tratamiento.

Diseño de un grupo, pretest y posttest

El agregado de un pretest al diseño anterior mejora la situación pero, aún así, sigue teniendo serias dificultades. En este caso tenemos el diseño de *un grupo, con pretest y posttest* (Diagrama 3.2). Veamos algunos de los factores que amenazan la validez, comenzando por la *historia*. Entre t_2 y t_4 pueden haber ocurrido algunos acontecimientos que provocaron que O_2 sea mayor que O_1 , con lo que no podemos atribuirle a X esta diferencia. Por ejemplo, el investigador quiere averiguar si una forma nueva de enseñar (X_1) representa una buena alternativa a la que usaba hasta ese momento (X_0). Realiza un experimento en el que el contenido a enseñar es *la inflación*. Supongamos que luego de administrado el pretest (entre O_1 y O_2), se produce una crisis económica y este tema estalla en los medios de comunicación, es posible que los alumnos aprendan por esto último y no por la enseñanza recibida.

Diagrama 3.2. Diseño de un grupo, con pretest y posttest

$$\begin{array}{cccc} t_1 & t_2 & t_3 & t_4 \\ NG & O_1 & X_1 & O_2 \end{array}$$

Dentro de la *historia*, incluimos los acontecimientos que se desarrollan en forma paralela al experimento y que influyen en la variable dependiente en el mismo sentido que el hipotetizado para la variable independiente. Es más probable que actúe cuanto más tiempo pase entre el pretest y el posttest. Considerando dos casos extremos, si se trata de un experimento que se lleva a cabo en una sola clase, es menos probable que actúe que si éste dura todo un año. En el ejemplo de Vidal-Abarca, (2000) visto más arriba (Capítulo 1), si se llevara a cabo en una sola sesión aunque sin el posttest diferido, es menos probable que la historia represente una amenaza. En este caso, podemos suponer que son escasas las posibilidades de que existan sucesos paralelos al experimento que hagan que los alumnos sometidos a un tratamiento determinado modifiquen su comprensión lectora, aunque, como dije antes, es fundamentalmente el conocimiento del contexto y en general de las condiciones en que se llevó a cabo el experimento lo que nos puede ayudar a descartar este factor (como cualquier otro).

De todos modos, es difícil que otras amenazas de validez sean controladas. Por ejemplo, la *maduración*. A veces se dan cambios en los sujetos por el solo paso del tiempo (que tienen sus causas pero que no conocemos o no podemos prever) y entonces este factor representa una amenaza. Se trata del concepto metodológico de *maduración*, pues en él se incluyen todos los cambios internos posibles de los sujetos participantes en el estudio mientras éste se desarrolla, por ejemplo, por cansancio.

Supongamos que queremos probar el efecto de un tratamiento psicoterapéutico para tratar determinadas neurosis y usamos este diseño. Se le hace un diagnóstico a los sujetos participantes antes de iniciar la terapia, tras lo cual ésta comienza. Luego de 6 meses (tiempo en que se supone que dicho tratamiento dará resultados), se vuelve hacer un diagnóstico y se constata que efectivamente, los sujetos mejoraron significativamente su salud mental. Aquí, si bien se comprobó que efectivamente hubo un cambio, no podemos estar seguros que el mismo no

se deba a otros factores, por ejemplo, que los sujetos hayan madurado por el paso del tiempo.

También puede suceder que la *aplicación del pretest* haya producido una modificación en la variable dependiente. Esto es así, pues el solo hecho de responder a una prueba modifica a los sujetos, por ejemplo, haciéndole tomar conciencia de ciertos aspectos en los que tal vez nunca hubieran pensado. Así, si le aplicamos el diagnóstico al grupo y sin mediar el tratamiento, luego de un tiempo volvemos a implementarlo, es muy probable que en la segunda observación se registren mejores resultados. Este factor se llama *aplicación de pruebas* (o test).

También podrían ocurrir cambios de quienes realizan los diagnósticos, en el modo de implementar las pruebas o en el de corregir las respuestas de los sujetos. En el ejemplo recién mencionado (se quería probar el efecto de un tratamiento psicoterapéutico), si se toman entrevistas abiertas para indagar los síntomas que presentan los sujetos, es posible que no se hagan las mismas preguntas en el pretest y en el postest. Aún haciéndolas, puede que luego las respuestas no se clasifiquen con un criterio similar al de la primera prueba. Si ésta es menos exigente que la primera, puede hallarse una diferencia entre O_2 y O_1 que no se debe a la aplicación del tratamiento. Esta amenaza se llama *instrumentación*.

La *regresión* es un fenómeno que se descubrió a partir de investigaciones empíricas. Alude al hecho de que aquellos sujetos que registraron puntajes extremos (muy altos o muy bajos) en una primera prueba, obtendrán una puntuación más próxima hacia la media en una segunda medición, aun si no fueron expuestos a ningún tratamiento experimental. Por esto, si los grupos no se formaron al azar, puede suceder que el experimental esté constituido mayoritariamente por sujetos que tienen un nivel inicial muy bajo (respecto de la variable dependiente) con lo que en el postest mostrarán, como norma general, registros más altos por el fenómeno de la *regresión*.

Para sintetizar lo dicho sobre este diseño, recordemos lo que Shadish et al. (2002) sostienen sobre él: “los científicos sociales (...) rara vez serán capaces de construir conocimiento causal confiable con el simple diseño pretest-posttest, a menos que los resultados hayan mejorado claramente [entre el pre y el posttest] y el intervalo entre las pruebas preliminares y posteriores sea corto” (Shadish et al. 2002: 110).

Diseño de un grupo, doble pretest y posttest

Al diseño de un grupo, pretest y posttest se le pueden hacer algunos agregados o modificaciones que lo mejoran, algunas de ellas son: 1) la inclusión de doble pretest; 2) el uso de una variable dependiente no equivalente; 3) la repetición en tiempos posteriores de las observaciones pero sin tratamiento (tratamiento removido) y; 4) esta situación con un tratamiento entre las dos últimas observaciones. Veámoslas en ese orden.

La inclusión de un doble pretest (Diseño de un grupo, posttest y doble pretest, Diagrama 3.3) reduce la amenaza de la maduración, pues si ésta está influyendo en la variable dependiente, sería esperable que ocurra entre O_1 y O_2 , es decir, si no hay diferencia entre estas dos observaciones, tenemos alguna estimación de su influencia. Y lo mismo podemos decir de otros factores como la historia, la aplicación de pruebas, la instrumentación o la regresión. En otras palabras, si estos factores estuvieran actuando, sería esperable que esto se dé entre O_1 y O_2 , además de observarse entre O_2 y O_3 . Es decir, que el investigador trata de reducir algunas amenazas agregando mediciones de la variable dependiente que le permitan estimar si es probable que algunos factores de validez influyan en la variable dependiente. De todos modos, igualmente debe ser muy cuidadoso.

Diagrama 3.3. Diseño de un grupo con doble pretest y posttest

t_1	t_2	t_3	t_4	t_5
NG_1	O_1	O_2	X_1	O_3

Diseño con pretest y posttest con variable dependiente no equivalente^{*32}

En este caso se mide la variable dependiente antes y después del tratamiento (O_{1y} y O_{2y} respectivamente), pero también se mide otra u otras variables diferentes antes del tratamiento, representadas por O_{1z} $O_{1a} \dots O_{1n}$ y después del mismo (representadas por O_{2z} $O_{2a} \dots O_{2n}$ ver

³² Los apartados señalados con un asterisco pueden evitarse sin perder continuidad en la lectura.

Diagrama 3.4). La primera como es lógico se espera que aumente en la segunda medición, pero la variable no equivalente no se espera que cambie (por el conocimiento que tenemos de la misma) como resultado del tratamiento. En el ejemplo planteado más arriba, el docente quiere saber si una forma de enseñanza es efectiva y para eso hace un experimento con el concepto de inflación. En el pretest observa lo que saben los alumnos sobre inflación y sobre interés. Luego hace lo propio en el postest. Si, por ejemplo, en ambos conceptos se produce una ganancia entre el pretest y postest, es posible que la maduración u otros factores hayan actuado.

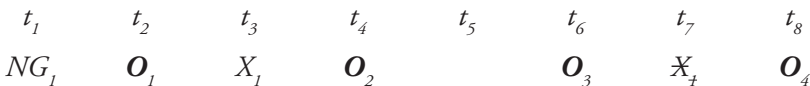
Diagrama 3.4. Diseño con pretest, postest y variable dependiente no equivalente



Diseño con tratamiento removido*

Shadish et al. (2002) usan la X tachada ($\bar{X}$) para indicar que el tratamiento es removido. Se dan en primer lugar los tres momentos del diseño con pretest y postest de un grupo, luego le sigue un tiempo sin nada, y finalmente dos nuevas observaciones pero sin tratamiento (Diagrama 3.5). Estrictamente hablando significa lo mismo que un espacio en blanco (o que X_0), pero con este tachado los autores quisieron representar que el tratamiento antes existió y ahora se eliminó.

Diagrama 3.5. Diseño con tratamiento removido



Es esperable que la variable dependiente suba o baje, según se dé la presencia o ausencia del tratamiento. O sea que se espera que entre O_1

y O_2 se observe un aumento y que entre O_3 y O_4 esto no ocurra o se dé una disminución.

Este diseño es más aconsejable cuando se supone que el efecto de X no perdura en el tiempo, o en todo caso debe tenerse presente esta posibilidad, ya que el efecto del primer tratamiento puede confundirse con el momento en que éste es removido.

Diseño de un grupo con tratamiento repetido*

El diseño con tratamiento repetido es similar al anterior. También tiene cuatro observaciones (Diagrama 3.6), entre O_1 y O_2 , se da el tratamiento, entre O_2 y O_3 éste no se aplica ($\times$), y luego entre O_3 y O_4 esto ocurre nuevamente. Queda claro en cuáles son las observaciones en las que se espera que la variable dependiente aumente y en cuales disminuya o mantenga su valor. Es más difícil que las amenazas se hagan efectivas, pues si estuvieran presentes en los momentos en que se aplica el tratamiento, también debieran estarlo cuando éste es removido.

Diagrama 3.6. Diseño con tratamiento repetido

t_1	t_2	t_3	t_4	t_5	t_6	t_7	t_8
$N G_1$	O_1	X_1	O_2	$\times$	O_3	X_1	O_4

Diseños de una serie cronológica

Podemos implementar diseños de una o más series cronológicas, en este apartado veremos el primer caso y más adelante con dos de ellas. Utiliza un solo grupo, es decir, se basa en comparar a los mismos sujetos antes del tratamiento y después de él (Diagrama 3.7). Pero, a diferencia del *diseño de un grupo con pretest y posttest*, se realizan varias mediciones de la variable dependiente, que pueden registrar cambios en tres dimensiones: “la forma del efecto (el nivel, la pendiente, la varianza y la ciclicidad), su permanencia (continuo o discontinuo), y su inmediatez (inmediato o diferido)” (Shadish et al. 2002: 173).

Los diseños de series cronológicas intentan reducir las amenazas de validez con el agregado de mediciones antes y después del tratamiento. En efecto, si algún factor estuviera actuando entre O_4 y O_5 (es decir, antes y después del tratamiento) también es muy probable que lo haga entre O_1 y O_2 , entre ésta y O_3 y a su vez entre O_3 y O_4 , entre O_5 y O_6 , entre ésta y O_7 y entre esta última y O_8 . Es decir que tenemos muchas posibilidades de advertir si algún factor distinto de X estuviera produciendo la diferencia entre O_4 y O_5 , lo que no sucede con el diseño de un grupo con pre y posttest.

Diagrama 3.7. Diseño de una serie cronológica

t_1	t_2	t_3	t_4	t_5	t_6	t_7	t_8	t_9	t_{10}
NG_1	O_1	O_2	O_3	O_4	X_1	O_5	O_6	O_7	O_8

Investigación explicativa causal con un solo sujeto ($N=1$)

Los estudios de sujeto único, o también conocidos como diseño $N=1$, además de ser interesantes por sí mismos, nos servirán para discutir algunos aspectos relativos a otros tipos de investigación. *A veces, suele denominarse del mismo modo al estudio con $N=1$, al estudio de casos y a la encuesta.* Sin embargo, si bien tienen algunos elementos en común, también es cierto que hay otros que los diferencian de un modo claro, por lo que *es conveniente no confundirlos.*

La encuesta puede ser vista como un método de investigación, como lo hacen Cohen y Manion (1990), o como una técnica de recolección de datos. En esta última posición podemos ubicar a Ander-Egg (2003), quien la llama “entrevista extensiva”. En este trabajo la concibo como método y puede responder a un estudio no experimental o descriptivo, según los objetivos que se propone el investigador: en el primer caso pone a prueba alguna hipótesis (con las limitaciones que veremos más adelante), mientras que en el segundo esto no sucede.

Los *estudios de casos*, aunque se los asocia a la investigación cualitativa, podrían incluir datos cuantitativos; estudian en profundidad al o los casos seleccionados como totalidades, se analizan varios aspectos al mismo tiempo, es decir, asumen un enfoque holístico del fenómeno en

estudio. En cambio, el *experimento* $N=1$ se usa para contrastar hipótesis causales, con utilización de datos cuantitativos y, a veces, test estadísticos (pruebas de significación) propios de la tradición cuantitativa. Consideran una sola variable independiente y una dependiente (o unas pocas), pero se limitan sólo a ellas, a diferencia del estudio de casos que puede ser descriptivo, de tipo cualitativo y para algunos de contrastación de hipótesis.

La *encuesta* a veces se puede considerar dentro de los estudios de casos, en el sentido de que su finalidad es describir una población que puede ser considerada como un caso. Pero se diferencian, porque la primera en general es extensa en cuanto a la población que abarca, aunque no profunda (realiza preguntas predeterminadas y generalmente cerradas, sobre cada aspecto o dimensión que pretende analizar), mientras que en el estudio de casos, generalmente se analizan en profundidad casos, siendo más importantes los registros narrativos antes que los datos cuantitativos (aunque como dije recién podemos pensar en estudios de casos de tipo cuantitativo).

El tipo de estudio $N=1$ más elemental presenta dos series en el momento empírico (Diagrama 3.8), una para establecer la línea base (X_0O_1 , X_0O_2 , X_0O_3 y X_0O_4 ; serie A), y otra con intervención (X_1O_5 , X_1O_6 , X_1O_7 y X_1O_8 ; serie B).

Diagrama 3.8. Diseño $N=1$ (AB), con dos series de cuatro observaciones cada una

	t_1	t_2	t_3	t_4	t_5	t_6	t_7	t_8	t_9
$N=1$	X_0O_1	X_0O_2	X_0O_3	X_0O_4	X_1O_5	X_1O_6	X_1O_7	X_1O_8	

Así, en *primer* lugar, se selecciona un sujeto que puede ser un alumno de un colegio, un paciente que asiste a tratamiento, un docente, un directivo, etc.

En *segundo* lugar, antes de la intervención, se lleva a cabo un conjunto de observaciones, que se denominan *serie o fase*, cuyo objetivo es establecer la línea base, registrando la conducta que se quiere modificar (variable dependiente) en varios momentos. Cuanto más grande es el

número de observaciones, mayor será su estabilidad. Esta *línea de base*, nos permite conocer qué valor asume la variable dependiente sin intervención o en general, sin tratamiento (nos da una estimación de ella, por ejemplo, tomando el promedio de todas las observaciones de la serie) y nos posibilita conocer la dispersión que tiene la misma.

En *tercer* lugar, se realiza la intervención, y durante este período también medimos la variable dependiente. Es conveniente concretar un número de observaciones al menos tan grande como en el período anterior. Las fases se pueden repetir, lo que nos da distintos diseños para este tipo de estudios.

Un ejemplo puede ser el siguiente. Se quiere investigar si cierto modo de organizar las actividades de enseñanza (con clases de corta duración y períodos de distracción entre cada actividad llevada a cabo en el aula) disminuye la conducta hiperactiva, ante lo cual se selecciona un niño que manifiesta esta característica y se lo observa durante cinco días sin realizar ningún cambio. En la semana siguiente, se establece esta nueva forma de enseñanza, también se registra la cantidad de veces que manifiesta signos de hiperactividad y, finalmente, se comparan las medidas de cada una de las series desarrolladas.

Podemos repetir las series obteniendo otros diseños $N=1$. El que se muestra en el Diagrama 3.9, tiene dos fases sin intervención (series *A*) y dos con intervención (series *B*). Siempre se comparan los valores que asume la variable dependiente en cada serie: con tratamiento o sin él³³. También se los suele llamar a estos diseños *AB* y *ABAB*, respectivamente, y en general se usa la “*A*” para representar cada momento con ausencia de tratamiento, y la “*B*”, cuando éste está presente.

Existen dudas sobre la interpretabilidad (validez interna) de los diseños $N=1$, por el hecho de trabajar con un solo sujeto. Comparándolos con los estudios anteriores, que se basan en estudiar a grupos de sujetos, León y Montero (1993) nos explican con toda claridad los porqués. Podemos sintetizar los fundamentos de los estudios con una muestra representativa, diciendo que “suponen una lucha constante contra las diferencias individuales” (: 226).

33 A veces hago la descripción admitiendo que los diseños tienen una variable dependiente, pero estos estudios pueden tener más de una, como todos los que se mencionan aquí.

Diagrama 3.9. Diseño N=1 (ABAB), con cuatro series de tres observaciones

Primera serie (A)				Segunda serie (B)			Tercera serie (A)			Cuarta serie (B)		
t_1	t_2	t_3	t_4	t_5	t_6	t_7	t_8	t_9	t_{10}	t_{11}	t_{12}	t_{13}
$N=1$	X_0	O_1	X_0	O_2	X_0	O_3	X_1	O_4	X_1	O_5	X_1	O_6
	X_0	O_7	X_0	O_8	X_0	O_9	X_1	O_{10}	X_1	O_{11}	X_1	O_{12}

Estas diferencias son tratadas como errores, que se disimulan en la media (o cualquier otra medida estadística). Por ello se dice que estudian a un sujeto inexistente: el sujeto medio. En cambio, el experimento $N=1$ tiene una lógica completamente diferente, centrada en el individuo: “si un fenómeno es lo suficientemente básico se tendrá que producir de forma similar en el resto de los individuos” (León y Montero, 1993: 227).

El experimento con sujeto único,

“únicamente es capaz de proporcionar una técnica experimental para evaluar intervenciones para el sujeto individual, [... además presenta] ambigüedades introducidas por las tendencias y variaciones en los datos de la fase de línea de base y con la generalidad de los resultados de la investigación” (Cohen y Manion, 1990: 266).

León y Montero (1993) señalan con mayor precisión en qué circunstancias se presentan las dificultades de interpretación (es decir, de validez interna) de los experimentos $N=1$: cuando no existe una diferencia clara entre los momentos en que el sujeto se expone al tratamiento experimental y cuando esto no sucede, cuando la variabilidad al interior de las series de observaciones es amplia o cuando los patrones de eficacia no están claramente establecidos.

Los estudios $N=1$, son más fuertes cuando se cumplen los dos supuestos siguientes: 1) El efecto del tratamiento experimental se observa únicamente mientras se realiza (es decir, no es prolongado) y, 2) este efecto se puede observar en un período corto de tiempo.

Si no se da el primer supuesto, sólo podemos usar el *diseño con dos series (AB)*, por el problema de la interferencia entre las distintas series (si usamos el *ABAB* por ejemplo, como el efecto del tratamiento de la

primera serie perduraría aún después de interrumpirlo, en la segunda sin tratamiento, tendríamos valores en la variable dependiente más altos que los reales para este caso). Si no se cumple el segundo, ocurre lo contrario: es probable que en la segunda serie el efecto resulte menor al real, por no presentarse (o no hacerlo con toda su intensidad) en los pocos días que ella dura.

Diseños cuasiexperimentales con grupo de control

Diseños sin pretest

Diseño de dos grupos y solo posttest

Este diseño surge de agregar un grupo al diseño de un grupo y solo posttest (Diagrama 3.10). Como los sujetos no se incluyeron en el grupo experimental y en el de control al azar, puede haber dificultades con la *distribución*. Más adelante veremos que existen formas de controlar (reducir su amenaza) este factor cuando la asignación a los distintos grupos no se realizó al azar. El diseño de dos grupos y solo posttest también deja dudas con la *mortalidad* que, de un modo general, se refiere al hecho de que se pierden datos sobre la variable dependiente en las mediciones posteriores a la aplicación de los tratamientos. Por ejemplo, se realiza un estudio para conocer si el uso de un determinado programa informático favorece el desarrollo del pensamiento matemático y se hace el experimento fuera del horario de clase con el *diseño de un grupo, con pretest y posttest*; en este caso es muy probable que abandonen quienes tienen menos interés en el tema, por lo que la segunda observación puede dar valores más altos debido a que en el grupo quedan los más motivados para aprender y no por lo que les produjo el software educativo. Cuando el experimento se hace con dos o más grupos, puede suceder que abandonen el experimento sujetos que tienen características diferentes en un grupo respecto del otro. Por ejemplo, si en el grupo de control dejan aquellos que están muy motivados y, recíprocamente, en el grupo experimental lo hacen quienes tienen poco interés en el tema estudiado, esto provocará una diferencia en los posttest que no fue producida por el tratamiento experimental (en este caso, quedarán en

el grupo experimental sujetos más motivados que en el de control y esto producirá una diferencia en los posttest que se confundirá con la acción de la variable independiente).

Por último, la *interacción* alude al hecho de que algunos de los factores que ya vimos, al combinarse entre sí, a veces producen un nuevo efecto sobre la variable dependiente. La falta de distribución al azar, de por sí, puede afectar la validez interna; pero aún estando controlada, al unirse con la maduración, la historia u otras condiciones, podría provocar un aumento de la variable dependiente, que se confundiría con la variable independiente, con lo que la variación observada entre grupos (o entre el pretest y el posttest aunque no es el caso que estoy tratando) no se podría atribuir legítimamente a los tratamientos. Todos estos problemas también se dan si la variable independiente tiene más de dos valores o categorías y por lo tanto, usamos más de dos grupos, con uno por cada uno de estos valores o categorías.

Diagrama 3.10 Diseño de dos grupos naturales, con posttest únicamente

t_1	t_2	t_3
NG_1	X_0	O_1
NG_2	X_1	O_2

*Diseños con grupos duplicados y sólo posttest**

Este diseño consiste en repetir el diseño de dos grupos naturales y sólo posttest (Diagrama 3.11). Supongamos que la regresión haya actuado en el grupo 1 pues éste tenía un nivel inicial muy bajo, pero hemos repetido tanto el tratamiento como los posttest en otros dos grupos, en este caso tenemos más posibilidades de advertir si este factor influyó, pues sería demasiada casualidad que los dos grupos experimentales tuvieran un nivel bajo antes de los tratamientos.

Diseños de cohorte

En los diseños de cohorte, el grupo experimental recibe el tratamiento en un tiempo posterior al grupo de control y si es de grupos repetidos esto se reitera luego en otros dos grupos (Diagramas 3.12 y 3.13).

Diagrama 3.11 Diseño con grupos duplicados y sólo posttest³⁴

t_1	t_2	t_3
NG_1	X_0	O_1
NG_2	X_1	O_2
NG_3	X_0	O_3
NG_4	X_1	O_4

Diagrama 3.12 Diseño de cohorte

t_1	t_2	t_3	t_4	t_5
NG_1	X_0	O_1		
NG_2			X_1	O_2

Diagrama 3.13 Diseño de cohorte con grupos repetidos³⁵

t_1	t_2	t_3	t_4	t_5	t_6	t_7	t_8	t_9
NG_1	X_0	O_1						
NG_2			X_1	O_2				
NG_3					X_0	O_3		
NG_4							X_1	O_4

Los diseños cuasiexperimentales de dos o más grupos, con pretest y posttest

34 Tomado de León y Montero (2003), aunque con la simbología usada aquí.

35 Tomado de León y Montero (2003), aunque con la simbología usada aquí.

Algunos diseños cuasiexperimentales tratan de reducir las amenazas de validez de modos diferentes, como por ejemplo, a través de un pretest, en los que cuentan con dos o más grupos y un posttest; y con la aplicación de todos los tratamientos a todos los grupos en los compensados. Comencemos por los primeros.

Obsérvese que este diseño guarda una similitud en su forma con el análogo riguroso (de dos grupos, pretest y posttest). Aunque a simple vista pareciera que la diferencia entre ellos es mínima, en realidad tiene su importancia, pues en el que trato aquí, los sujetos no se distribuyen al azar en los distintos grupos, lo que hace que la adopción de la hipótesis de la equivalencia entre éstos sea más difícil de aceptar (Diagrama 3.14). No obstante, ésta resultará más plausible si O_{11} y O_{12} , arrojan resultados semejantes, es decir cuando los pretest resulten similares.

Otra forma de ver la validez se relaciona a los supuestos que implica cada diseño, es decir, los enunciados que debemos suponer sin certeza (se deben chequear por otros medios o asumirse corriendo el riesgo de equivocarse). Uno de ellos es la equivalencia de los grupos antes de la aplicación de los tratamientos. Este diseño, al agregar una observación a cada grupo antes de la aplicación de X (los tratamientos), permite advertir si ellos son similares o no en la variable dependiente cuando aún no se intentó modificar la variable independiente.

De todos modos, tiene menos posibilidades de controlar los factores de validez que si fuera riguroso, pues no produce una neutralización por aleatorización, sino que mediante el agregado del pretest pretende estimar la equivalencia entre los grupos previa a la aplicación de X (control por agregado de mediciones de la variable dependiente). Además, puede realizarse con tres o más grupos, lo cual es necesario si la variable independiente tiene más de tres categorías o valores.

Diagrama 3.14 Diseño de dos grupos naturales, pretest y posttest

t_1	t_2	t_3	t_4
$N G_1$	O_{11}	X_0	O_{21}
$N G_2$	O_{12}	X_1	O_{22}

Diseños de dos o más series cronológicas*

El *diseño de dos series cronológicas* representa una variación del que cuenta con una sola y vimos más arriba. Agrega un grupo que no se expone al tratamiento experimental, lo que da mayores posibilidades de control. Lo ideal es que el grupo agregado sea lo más similar que se pueda al grupo experimental (Diagrama 3.15).

Diagrama 3.15. Diseño de dos series cronológicas

t_1	t_2	t_3	t_4	t_5	t_6	t_7	t_8	t_9	t_{10}
NG_1	O_{11}	O_{21}	O_{31}	O_{41}	X_0	O_{51}	O_{61}	O_{71}	O_{81}
NG_2	O_{12}	O_{22}	O_{32}	O_{42}	X_1	O_{52}	O_{62}	O_{72}	O_{82}

Diseños compensados

Este tipo de estudio se basa en exponer a un mismo grupo a varios tratamientos en forma sucesiva. El primero que presento es el más elemental y cuenta con dos grupos, que reciben sucesivamente los dos tratamientos (Diagrama 3.16).

Este diseño permite obtener luego una media para cada grupo y para cada tratamiento, lo que posibilita conocer si las diferencias entre las distintas observaciones se deben a unos u otros. La Tabla 3.1 muestra el modo de obtenerlas. O_{11} y O_{21} son las observaciones registradas cuando el grupo 1 es expuesto a X_0 y X_1 , si las sumamos y dividimos por 2 obtenemos la media de este grupo ante los dos tratamientos (M_1). Recíprocamente, si hacemos lo mismo con O_{12} y O_{22} , obtenemos la media para el grupo 2. Al mismo tiempo, M_{x0} , surge de promediar los valores obtenidos en los dos grupos con el mismo tratamiento (X_0) y recíprocamente, M_{x1} , de promediar las cifras registradas en los dos grupos con X_1 . Cuanto más se parezcan M_1 y M_2 , más podremos aceptar la hipótesis de la equivalencia preexperimental de los grupos y cuanto más se diferencien M_{x0} con M_{x1} , mayor será la influencia de X (estas medias se obtienen de tratar a los dos grupos con cada uno de los tratamientos).

Estos estudios también pueden implementarse con una variable independiente que cuente con más valores, lo que implica formar más grupos. Por ejemplo, si X supone tres posibilidades, tenemos el diseño de tres grupos y tres tratamientos (Diagrama 3.17).

Generalizando podemos decir que el número de grupos es igual al número de tratamientos (dentro de estos por supuesto debemos contar los placebos o los tratamientos nulos, si los hay).

Diagrama 3.16 Diseño compensado de dos grupos y dos tratamientos

t_1	t_2	t_3	t_4	t_5
$N G_1$	X_0	O_{11}	X_1	O_{21}
$N G_2$	X_1	O_{12}	X_0	O_{22}

Tabla. 3.1. Medias de grupos y tratamientos en el diseño compensado de dos grupos y dos tratamientos

$(O_{11} + O_{21}) / 2 = M_1$	Media del Grupo 1
$(O_{12} + O_{22}) / 2 = M_2$	Media del Grupo 2
$(O_{11} + O_{22}) / 2 = M_{x0}$	Media del tratamiento X_0
$(O_{12} + O_{21}) / 2 = M_{x1}$	Media del tratamiento X_1

Diagrama 3.17 Diseño compensado de tres grupos y tres tratamientos

t_1	t_2	t_3	t_4	t_5	t_6	t_7
$N G_1$	X_0	O_{11}	X_1	O_{21}	X_2	O_{31}
$N G_2$	X_1	O_{12}	X_2	O_{22}	X_0	O_{32}
$N G_3$	X_2	O_{13}	X_0	O_{23}	X_1	O_{33}

Ejercicios

1. Del Apéndice B, seleccioná el o los estudios cuasiexperimentales y realizá el diagrama del diseño.
2. Si tuvieraas que replicarlo, ¿qué tratarías de modificarle para que tengan más validez? Justificá tu respuesta.

3. Pensá en una investigación cuasiexperimental que podrías realizar en tu Trabajo Final (o en alguna otra exigencia académica) y describila en media página; realizá el diagrama de su diseño.
4. Modificalo para que se transforme en riguroso.

Capítulo 4

Diseños no experimentales

En este trabajo llamo *no experimentales* a los estudios en que se ponen a prueba hipótesis que afirman relaciones de causalidad entre variables, pero éstas sólo son observadas, sin que se las manipule como en los de corte experimental. Estrictamente, sólo permiten conocer en qué medida están vinculadas, aunque a veces el investigador, basándose en ciertos conocimientos previos a su trabajo, puede interpretar una asociación hallada en términos de causa y efecto. No deben ser visos sin embargo como inferiores a los experimentales, pues como ya dije en varias oportunidades, en determinadas situaciones podrían tener más validez que éstos, en función de las características específicas del estudio (Shadish et al. 2002).

Con estos diseños se puede analizar si dos o más variables están relacionadas de dos modos diferentes: 1) partiendo de una hipótesis o, 2) simplemente buscándola luego de recolectados los datos. El segundo caso tiene como finalidad el descubrimiento de una relación entre variables, trabajando en forma inductiva; en el primero se adopta al menos una hipótesis a priori y se intenta contrastarla, haciéndolo en forma hipotético deductiva.

Los estudios no experimentales también se llaman correlacionales, pues normalmente se utiliza el coeficiente de correlación como instru-

mento de análisis de datos. Éste básicamente nos informa sobre el grado en que se da la relación entre dos variables y sobre el sentido de la misma (recuérdese la cita de Mayntz, et. al. 1975, incluida en el Capítulo 1³⁶). Sin embargo, no hay que creer por ello que en la investigación experimental no se usan coeficientes de correlación, pues en ésta también podrían considerarse a la hora de analizar la información.

De cualquier manera, debemos tener presente que “una correlación no implica necesariamente una relación de causa a efecto entre dos factores” (Cohen y Manion, 1990: 213), pero la ausencia de correlación supone ausencia de causalidad. Ello es así, pues causalidad implica necesariamente correlación. Esto significa que *los estudios no experimentales nos permiten refutar hipótesis causales, aunque una corroboración debe interpretarse con más cuidado aún que en estudios cuasiexperimentales o experimentales rigurosos*³⁷.

Según Cohen y Manion (1990: 209) la investigación no experimental resulta adecuada en investigaciones exploratorias en campos en donde existe poca o ninguna investigación previa,

“cuando hay una necesidad de descubrir o aclarar relaciones [...] Es especialmente útil, en ese sentido, en los pasos iniciales de un proyecto donde tiene que cubrirse una cierta cantidad de trabajo básico de fundamentación para conseguir una cierta idea de la estructura de las relaciones” (Cohen y Manion, 1990: 209)³⁸

y también cuando el objetivo es realizar predicciones. Además, tienen una enorme ventaja sobre los estudios cuasiexperimentales y rigurosos, ya que permiten analizar simultáneamente las relaciones entre un gran número de variables, lo que los aproxima más a las situaciones reales. Esto los diferencia de los dos últimos, que sólo pueden considerar una sola causa (variable independiente) por experimento, o unas pocas

36 El sentido de la relación puede ser positivo o negativo, es decir, la relación puede ser directa o inversa. En el primer caso, al aumentar una variable, también lo hace la otra, y recíprocamente, al disminuir la primera, ocurre lo mismo con la segunda. En el segundo caso, cuando una variable aumenta, la otra disminuye y viceversa.

37 El razonamiento se basa en la regla lógica generalmente denominada *modus tollens*, que puede esquematizarse así: p implica q , y no- q , por lo tanto no- p . En este caso tenemos: si hay una relación de causalidad entre X e Y , entonces hay correlación, pero no hay correlación, por lo tanto, no hay causalidad. Esto a su vez, es equivalente a afirmar que de una hipótesis causal se deduce necesariamente una hipótesis correlacional y la inversa no es verdadera, pero la falsedad de una hipótesis correlacional implica la falsedad de una hipótesis causal.

38 Estos autores hablan de investigación correlacional en vez de no experimental.

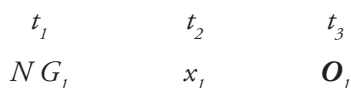
si se trata de un estudio factorial, en situaciones más artificiales, lo que supone menores posibilidades de generalizar.

Ya dije que los estudios no experimentales pueden ser muy importantes para descubrir relaciones causales que luego se contrasten con métodos que permitan un control más riguroso, pero incluso a veces son la única alternativa que tenemos, pues, existen variables que no se pueden manipular, por razones éticas, por razones económicas o simplemente porque no se conoce la forma de hacerlo. Por ejemplo, no podemos modificar la clase social de los alumnos para conocer qué influencia tiene ésta en la elección de una carrera universitaria, o no podemos extraerle el lóbulo frontal a un grupo de sujetos con la intención de indagar si éste influye en la capacidad de planificar o controlar la agresividad en los individuos estudiados.

Diseños no experimentales prospectivos

Un diseño no experimental prospectivo es aquel en que nos proponemos observar la o las variables independientes en primer lugar y luego la o las dependientes, refiriendo estas últimas a hechos que no han sucedido al momento de medir las primeras. El más elemental de todos cuenta con un solo grupo, lo que naturalmente no permite hacer comparaciones. Supongamos que tenemos en mente la siguiente hipótesis: la comprensión lectora influye en el rendimiento de los alumnos universitarios y que tomamos a todos los alumnos que tienen baja comprensión lectora y medimos su rendimiento. Si este también es bajo, ¿podemos concluir que las variables están asociadas? Podría suceder que los que tienen alta comprensión también muestren un bajo rendimiento (lo que se daría si las variables no están relacionadas), lo que no nos permite advertir el *Diseño no experimental de un grupo y sólo posttest* (Ver Diagrama 4.1).

Diagrama 4.1. Diseño prospectivo de un grupo y sólo posttest



Supongamos ahora que le tomamos a una muestra de alumnos una prueba de comprensión lectora al iniciar las clases o en las primeras semanas de cursado y conformamos dos grupos, por un lado, quienes manifestaron un alto nivel de comprensión y aquellos en que éste fue bajo. Después comparamos el rendimiento en ambos grupos³⁹.

Nótese que se midió la variable independiente antes que la dependiente, o antes de que los hechos que se describirían con la variable dependiente ocurrieran⁴⁰, por ello lo caracterizo como prospectivo. El Diagrama 4.2 corresponde al mismo y en él utilizo una *x* minúscula para señalar que la variable independiente no se manipuló.

Otro ejemplo puede ser el siguiente: queremos poner a prueba la hipótesis que afirma que “quienes han asistido a un colegio secundario con orientación humanística tendrán un mejor desempeño en la universidad en carreras de esta clase”. En este caso, podríamos conformar dos grupos: uno con quienes siguieron esta orientación y otro con quienes no lo hicieron. Finalmente, indagamos si hay diferencia entre ambos grupos en el rendimiento. Vemos que sólo es posible medir las variables, sin que el investigador decida quiénes van a seguir una carrera u otra, por lo que *no hay manipulación, lo que sucede en todos los estudios no experimentales*.

Estos diseños no experimentales pueden realizarse con un número indeterminado de grupos. Por ejemplo, si lo queremos concretar con tres de ellos, sería como se apreciaba en el Diagrama 4.3.

Diseño no experimental prospectivo de dos grupos, pretest y posttest

El primer diseño que vimos, en algunas situaciones particulares, se podría realizar con pretest, aunque es poco probable que se dé una situación para que podamos hacerlo. Por ejemplo, si estamos analizando la influencia de las estrategias de lectura en la comprensión lectora y sabemos que un determinado docente piensa implementar un curso

39 Este objetivo corresponde a uno de los desarrollados en un estudio que involucró a otros (Echevarría, 2008b). Los diseños en el sentido restringido en que los concibo aquí se pueden pensar por objetivos, incluso con la misma base de datos se pueden desarrollar diferentes diseños, por ejemplo descriptivos y no experimentales.

40 Suponiendo que la variable dependiente tiene dos categorías, aquí los hechos serían que los alumnos presentan alto rendimiento en tal o cual porcentaje, y bajo rendimiento en tal otro.

sobre ellas, mientras que en otro curso esto no ocurrirá, podríamos tomar a todos los alumnos una prueba de comprensión antes de que se inicie dicho curso y otra después, con lo que tendríamos un diseño con pretest y posttest, aunque sin manipular la variable independiente (Diagrama 4.4). Este ejemplo es de una situación ideal y como dije recién poco probable, no obstante, lo incluí para transmitir la idea de que el investigador a veces puede aprovechar situaciones específicas para lograr determinados objetivos si cuenta con un poco de sentido de la oportunidad e imaginación.

Diagrama 4.2 Diseño prospectivo de dos grupos y sólo posttest

t_1	t_2	t_3
$N G_1$	x_0	O_1
$N G_2$	x_1	O_2

Diagrama 4.3. Diseño prospectivo de tres grupos y sólo posttest

t_1	t_2	t_3
$N G_1$	x_0	O_1
$N G_2$	x_1	O_2
$N G_3$	x_2	O_3

Diagrama 4.4. Diseño no experimental de dos grupos, pretest y posttest

t_1	t_2	t_3	t_4
$N G_1$	O_{11}	x_0	O_{21}
$N G_2$	O_{12}	x_1	O_{22}

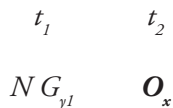
Las x minúsculas indican que la variable independiente no se manipula, en algunas situaciones se observa al mismo tiempo que la dependiente pero se incluyen antes para representar el hecho de que se analiza el efecto como si se hubiera manipulado.

Diseños retrospectivos

Tal como ponen de manifiesto los análisis epistemológicos, siempre la causa ocurre antes que el efecto. En el ejemplo que vimos (se estudiaba la relación entre la comprensión lectora y el rendimiento de los alumnos universitarios), como no puede ser de otro modo, se dio antes la variable independiente que la dependiente (los alumnos ya tenían determinada comprensión antes de presentar tal o cual rendimiento) y también sucede que se miden en ese orden, pero no siempre esto es así. Por ejemplo, a veces tomamos los sujetos en función de la variable dependiente, y luego observamos (o estimamos) que ocurrió con la independiente. O lo que es lo mismo, indagamos hechos que ya han ocurrido, buscando variables que los expliquen. Por eso también son llamados *ex post facto*, aunque algunos mencionan de este modo a todos los estudios no experimentales.

El más elemental de los diseños no experimentales retrospectivos consiste en tomar un grupo de sujetos que tiene una característica determinada que corresponde a una categoría de la variable dependiente, y luego se le mide la variable independiente (Diagrama 4.5). Está basado en una lógica muy simple aunque con limitaciones que hay que evitar si esto es posible. Por ejemplo, queremos explicar los problemas de aprendizaje en función de la motivación, entonces tomamos un grupo que los tiene y le tomamos un cuestionario sobre motivación. Si hallamos que el grupo no está motivado hacia el aprendizaje concluimos que la hipótesis se cumple. La dificultad que presenta este diseño es que no permite comparar con un grupo sin problemas de aprendizaje. Aquí lo denomino *Retrospectivo de un grupo*, pues se observa primero la variable dependiente y luego la independiente (a ésta se la busca retrospectivamente, o sea, luego de que se produjo el efecto). NG_{yI} significa el grupo que tiene la categoría o valor I en la variable y .

Diagrama 4.5. Diseño retrospectivo de un grupo



La situación mejora si podemos considerar más valores o categorías de la variable dependiente. Veamos un ejemplo. Cardozo y Alderete (2009), estudiaron si algunas variables, clasificadas en individuales y sociales, favorecen la resiliencia. Las variables independientes consideradas fueron: síntomas de salud mental, conducta de adaptación social, competencia social, autorregulación de habilidades, autoconcepto, inteligencia, eventos negativos de vida y soporte social. A través de la regresión múltiple logística llegaron a la conclusión de que la variable que mejor predice la resiliencia es el autoconcepto, siguiéndoles el soporte social y la autorregulación de habilidades.

“La resiliencia se evaluó administrando instrumentos que ponen en consideración los siguientes aspectos: síntomas de salud mental, adaptación social y competencias sociales. Para formar los grupos de resilientes y no resilientes se consideraron como puntos de corte los percentiles 25 y 75” (: 161). Así, se ve que formaron los dos grupos en relación a la variable dependiente (el efecto, que fue la resiliencia), considerando dos categorías (resiliente y no resiliente). Para establecerlas se basaron en los percentiles, tomando los casos más extremos (por encima del percentil 75 y por debajo del 25).

Llamo a este diseño *Retrospectivo de dos grupos* (ver Diagrama 4.6), ya que se observa la variable dependiente primero o al mismo tiempo que las independientes (aunque no por esto pensamos que éstas se dieron antes). En el gráfico, para reflejar este hecho, la dependiente aparece en el momento de conformar los grupos. Además, puede verse también que en las O se mide x , que es la variable independiente. En este caso, O_{1x} quiere decir observación uno de la variable x , se le pone un número para diferenciarla de O_{2x} , que es la observación de la variable x en el otro grupo, dándose ambas al mismo tiempo. En otros términos, $1x$ y $2x$ no deben interpretarse como las categorías de la variable independiente, pues en esa medición se observan todas las que incluye esta variable (el número está antes de la variable, mientras que cuando son categorías o valores de ella está después, como en los diseños que no son retrospectivos).

El agregado de un grupo, a modo de control, mejora la situación respecto del diseño retrospectivo de un grupo. Se parte de formar grupos en función de la variable dependiente y se mide la variable independiente. Los hechos que se quieren explicar ya ocurrieron cuando hacemos la investigación (la presencia o ausencia de resiliencia en el ejemplo), por

eso es retrospectivo. Como acabo de decir, en este caso, O_{1x} significa observación 1 de la variable x , mientras que NG_{y1} , grupo perteneciente a la categoría 1 de la variable Y . Se puede agregar cualquier número de grupos, lo que quise representar con el *Diseño Retrospectivo de n grupos*.

Diagrama 4.6. Diseño retrospectivo de dos grupos

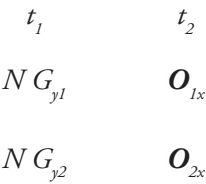
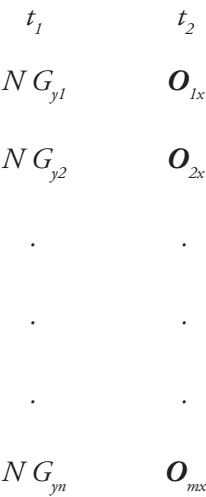


Diagrama 4.7. Diseño retrospectivo de n grupos



¿Grupos o situaciones? Diseño no experimental de una situación y sólo posttest

Hay que notar que en los diseños no experimentales, en muchas oportunidades los grupos sólo existen virtualmente, en el sentido de que es el investigador el que los considera de ese modo en el momen-

to de realizar el análisis de los datos. Por ejemplo, desarrollamos una encuesta en un colegio para conocer la relación que existe entre la profesión de los padres y el nivel de apoyo que estos les brindan a los alumnos en tareas extraescolares. En este caso, el investigador procesa la información como si hubieran existido grupos diferentes, cuando en realidad éstos nunca se conformaron físicamente, lo que sí sucede generalmente en los experimentales. Además, suele ocurrir que no conocemos a priori la cantidad de valores que asumirán las variables que consideraremos independientes, pues ignoramos sus distribuciones en la muestra o población a estudiar, con lo que tampoco podemos saber, antes de recolectar los datos, cómo se conformarán los grupos. Por esta razón, a veces se parte de estudiar una situación, lo que simbolizo con la letra “S” (Diagrama 4.8).

Una dificultad de este tipo de diseños es que la muestra podría no resultar adecuada para los objetivos propuestos (lo que puede suceder con otros no experimentales). En el ejemplo que acabamos de mencionar, si todos los padres (o la inmensa mayoría) tienen la misma profesión, no podríamos hacer las comparaciones necesarias entre ellos para analizar la relación que hay entre las variables de interés. Como norma general, todos los grupos que conformaremos con la o las variables independientes deben estar bien representados en cuanto al número de casos que contienen.

Diagrama 4.8. Diseño no experimental de una situación y sólo postest

t_1	t_2	t_3
S_1	$x_?$	O_1

Diseño no experimental de dos situaciones y sólo postest

También puede suceder que tomemos dos o más situaciones. En el caso hipotético mencionado recién (analizamos la relación entre la profesión y el apoyo en actividades extraescolares), en vez de un colegio, podríamos considerar dos, con lo que tenemos el *Diseño no experimental de dos situaciones y sólo postest* (Diagrama 4.9). Incluso el número de éstas podría ser mayor, lo que nos ofrece más garantías de que la muestra se adecuará a los objetivos de la investigación.

Diagrama 4.9. Diseño no experimental de dos situaciones y sólo posttest

t_1	t_2	t_3
S_1	$x_?$	O_1
S_2	$x_?$	O_2

Los diseños no experimentales suelen ser la única alternativa posible cuando tenemos muchas variables o cuando ellas cuentan con un número considerable de valores. Por ejemplo, si la variable independiente tuviera 10 valores, con un diseño experimental deberíamos conformar como mínimo igual cantidad de grupos, lo cual, por razones prácticas, no suele ser posible. Al optar por un estudio no experimental, si, además, no conocemos a priori cómo se distribuye la población según las variables independientes, no nos quedan más alternativas que optar entre los diseños que contemplan situaciones sin saber de antemano los grupos que se van a formar.

León y Montero expresan esta idea en los siguientes términos:

“Tanto en los diseños experimentales como en los cuasi-experimentales, los valores de la variable independiente eran elegidos por el investigador y no solían pasar de tres o cuatro (...). En los diseños «ex post facto» retrospectivos tanto los valores de la variable dependiente como los de las potenciales variables independientes vienen ya dados. El investigador elige los sujetos en función de la variable dependiente. Eso le garantiza la ocurrencia del fenómeno que quiere observar, sin embargo no sabe qué valores de las potenciales variables independientes van a aparecer en su muestra” (2004: 368).

Según ellos se debe aumentar lo más que se puede el tamaño y la representatividad de la muestra para tener las mayores garantías de que todos los valores de las variables independientes estén representados.

Pero si esto es así, la representatividad de la muestra tiene que ver tanto con la validez interna como externa. León y Montero (2004) sostienen que en los que ellos llaman los estudios de cuasi control, esto no sólo afecta a la validez externa, sino también a la interna. Además, sin contamos con una muestra aleatoria, cuanto más grande sea su tamaño, mayor será la validez externa.

Es decir que, si la muestra es representativa, podemos estar seguros de que la covariación observada en ella se da en la población, pero no necesariamente esta variación implica que exista una relación de causalidad entre las variables estudiadas.

Por ejemplo, queremos estudiar la relación entre el género (como causa o variable independiente) y el tipo de carrera que eligen los alumnos universitarios (como variable dependiente o efecto), formando grupos con la segunda (carreras humanísticas o relacionadas a las ciencias humanas, carreras relacionadas a las ciencias naturales y carreras relacionadas a las ciencias formales). En un estudio retrospectivo, trataremos de conformar la muestra de modo que haya una similar cantidad de sujetos de las tres carreras y luego observamos el género, mientras que si es prospectivo, seleccionamos sujetos tratando de que ambos géneros (femenino y masculino) queden representados en forma similar y luego observamos la carreras de los sujetos de cada uno de ellos.

Ejercicios:

1. Del Apéndice B, seleccioná el o los estudios no experimentales y realizá el diagrama del diseño.
2. ¿Podría indagarse el mismo objetivo en con otro tipo de diseño (por ejemplo, riguroso, cuasiexperimental)? Si considerás que la respuesta es afirmativa, describí en media página cómo lo harías.

Capítulo 5

Diseños factoriales

Los diseños experimentales se pueden implementar para una o más variables independientes y, en el último caso, se llaman *factoriales*. El Cuadro 5.1 muestra la clasificación, tanto de los rigurosos como de los cuasiexperimentales y no experimentales en función de la cantidad de variables independientes que consideran. Por razones didácticas los presento en este Capítulo destinado exclusivamente a ellos.

Cuadro 5.1. Diseños de contrastación de hipótesis

Experimentales rigurosos o estrictos	Simples (una variable independiente)
	Factoriales (más de una variable independiente)
Cuasiexperimentales	Simples (una variable independiente)
	Factoriales (más de una variable independiente)
No experimentales	Simples (una variable independiente)
	Factoriales (más de una variable independiente)

Los estudios factoriales pueden pensarse como una ampliación o generalización de los anteriores, pues se diferencian de ellos porque el investigador considera en el mismo estudio, más de una causa. Esto le permite analizar, por un lado, el efecto principal o global de cada una de las variables independientes y, por otro lado, la interacción que puede haber entre ellas. Como contrapartida, presentan la enorme dificultad de requerir una mayor cantidad de grupos, por lo que suelen ser difíciles de implementar⁴¹.

Al hablar de efecto *global*, me refiero al que una variable produce sobre otra sin incluir en el análisis una tercera, cuarta o más variables, mientras que la *interacción* es la influencia conjunta que dos o más de ellas ejercen sobre, por lo menos, una tercera. Supongamos que analizamos la relación que hay entre la *forma de enseñanza* (considerando dos alternativas: *directiva* y *no directiva*), *cociente intelectual* (*alto* y *bajo*) y *nivel de aprendizaje* (*distintos grado de logro de los objetivos*; ésta es la variable dependiente). Si la forma de enseñanza produce igual resultado para cualquier nivel intelectual, entonces no hay interacción. Pero si la enseñanza no directiva es más efectiva con el grupo de nivel intelectual alto que en el bajo y viceversa⁴², entonces hay interacción entre las dos variables independientes (en algún sentido interaccionan para producir un efecto que no puede explicarse por la suma de los efectos globales de las dos variables independientes).

Lo anterior es lo mismo que decir que la *interacción* existe cuando la acción de una variable sobre otra es diferente en ciertos grupos conformados en función de una tercera. En el ejemplo hipotético anterior, el efecto de la forma de enseñanza es específico (particular) para cada nivel intelectual y por ello, el último modifica la acción que produce la primera (interacciona con ella). Puede verse también en el ejemplo, por qué para analizar la interacción entre variables es necesario un diseño factorial.

Aunque el problema de la validez no ha recibido para los estudios factoriales el mismo tratamiento que para los de una variable, podemos decir con toda seguridad que *pueden ser con control riguroso o sin él, según cumplan o no con los requisitos de aleatorización* vistos al tratarlos.

41 El número de grupos que debemos formar es igual al producto del número de valores o categorías de cada variable independiente.

42 Lo mismo si sucede al revés.

En lo que resta de este capítulo, veremos algunos de los factoriales que podríamos implementar.

Diseños factoriales con control riguroso

El más elemental de todos los diseños factoriales tiene dos variables independientes (que por convención llamo W y X) y ambas son dicotómicas (tienen sólo dos valores W_0 , W_1 para W y X_0 , X_1 para X , ver Diagrama 5.1). En este caso, si queremos realizar un experimento riguroso, debemos formar cuatro grupos al azar, exponiendo a un grupo a W_0 , X_0 , otro a W_0 , X_1 , haciendo lo propio para W_1 , X_0 y W_1 , X_1 , con los dos restantes. Por ejemplo, supongamos que queremos poner a prueba la hipótesis que afirma que “El trabajo en grupo favorece el aprendizaje si la duración de la clase es suficiente como para que haya una buena interacción entre los alumnos; en cambio, si la enseñanza se distribuye en secciones de corta duración, es preferible el trabajo individual”.

Consideremos dos valores para la variable *distribución del tiempo* (dos sesiones de 35 y un recreo de 10, y 80 minutos respectivamente, que llamaremos W_0 y W_1). Utilizo hasta la letra X para mencionar a las variables independientes, por ejemplo, V , W , X . En este caso, la variable *forma de enseñanza* estaría representada por X (X_0 = *enseñanza individual* y X_1 = *enseñanza grupal*). Entonces al grupo 1 se le enseñará con dos sesiones de 35 minutos y con trabajo individual; al grupo 2, con dos sesiones de 35 minutos y trabajo grupal; en el grupo 3, las clases serán de 80 minutos ininterrumpidos en forma individual, y el grupo 4, con la misma forma pero los alumnos aprenderán en grupo⁴³. O sea:

W_0 : dos sesiones de 35 minutos y un recreo de 10,

W_1 : 80 minutos ininterrumpidos,

X_0 : *enseñanza individual*,

X_1 : *enseñanza grupal*.

43 No hay que confundir este diseño con el de *cuatro grupos de Solomon*, pues aquí éstos se forman por el hecho de tener dos variables independientes, mientras que en aquél, con la finalidad de repetir el *diseño de dos grupos, con distribución al azar y posttest solamente*, pero con la realización del pretest en los dos grupos agregados.

Este mismo experimento podría llevarse a cabo con la realización de un pretest, como se aprecia en el Diagrama 5.2 (*diseño factorial de control riguroso, 2 x 2, pretest y posttest*).

Las variables podrían tener más de dos valores. Por ejemplo, si X tuviera tres, estaríamos ante un *Diseño factorial de control riguroso, 3 x 2, pretest y posttest* (en el ejemplo antes citado, la forma de enseñanza podría tener como opciones: tres sesiones de 20 minutos y dos recreos de 10, dos sesiones de 35 y un recreo de 10, y una sesión de 80 minutos, (Diagrama 5.3).

Diagrama 5.1. Diseño factorial de control riguroso, 2 x 2, y sólo posttest

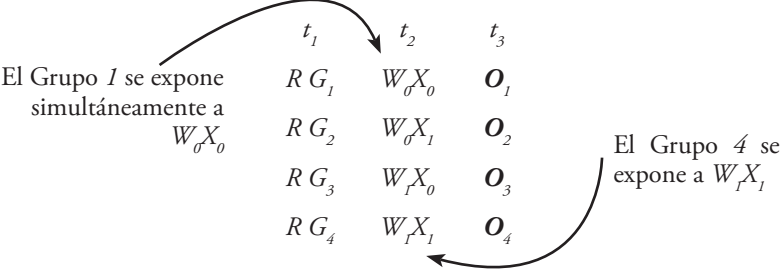


Diagrama 5.2. Diseño factorial de control riguroso, 2 x 2, con pretest y posttest

t_1	t_2	t_3	t_4
$R G_1$	O_{11}	$W_0 X_0$	O_{21}
$R G_2$	O_{12}	$W_0 X_1$	O_{22}
$R G_3$	O_{13}	$W_1 X_0$	O_{23}
$R G_4$	O_{14}	$W_1 X_1$	O_{24}

Diagrama 5.3. Diseño factorial de control riguroso, 3 x 2, con pretest y posttest

t_1	t_2	t_3	t_4
$R G_1$	O_{11}	$W_0 X_0$	O_{21}
$R G_2$	O_{12}	$W_0 X_1$	O_{22}
$R G_3$	O_{13}	$W_0 X_2$	O_{23}
$R G_4$	O_{14}	$W_1 X_0$	O_{24}
$R G_5$	O_{15}	$W_1 X_1$	O_{25}
$R G_6$	O_{15}	$W_1 X_2$	O_{26}

*Diseños factoriales cuasiexperimentales**

Así como se realizan diseños factoriales rigurosos, también podemos pensar en otros cuasiexperimentales. Aunque hay que hacer una aclaración. Si consideramos que son factoriales sólo aquellos diseños a los que se les puede aplicar el método de análisis factorial de variancia, los cuasiexperimentales no lo serían (por carecer del requisito de la aleatoriedad). Hernández Pina (1998), al tratar los diseños factoriales, habla de ANOVA de dos y de tres factores (aunque aclara que existen otros, no menciona cuáles son, por lo que no sabemos si para ella todos los factoriales son de tipo ANOVA o no). Al referirse al de dos factores, sostiene: “este diseño exige que los sujetos hayan sido distribuidos al azar a cada uno de los niveles de tratamiento” (: 105), lo cual es cierto si queremos aplicar este procedimiento de análisis estadístico. Pero aquí estoy usando el término *factorial en sentido amplio, para referirme a aquellos estudios que consideran más de una variable independiente, incluso si la distribución no se realiza al azar*.

En este caso, son más limitados en cuanto a las pruebas de significación que se pueden implementar, por carecer de distribución al azar, pero a veces no tenemos otra alternativa. Por ejemplo, si en lugar de asignar los sujetos al azar en el diseño de cuatro grupos con pretest y posttest tomamos grupos naturales, obtenemos su equivalente cuasiexperimental (Diagrama 5.4).

Diagrama 5.4. Diseño factorial sin distribución al azar, 2 x 2, con pretest y posttest

t_1	t_2	t_3	t_4
NG_1	O_{11}	$W_{00}X_0$	O_{21}
NG_2	O_{12}	$W_{01}X_1$	O_{22}
NG_3	O_{13}	$W_{10}X_0$	O_{23}
NG_4	O_{14}	$W_{11}X_1$	O_{24}

Si tenemos tres variables independientes dicotómicas, el número de grupos a formar será 8 (ver Diagrama 5.5), lo que torna extremadamente difícil su implementación. Éste surge de *multiplicar entre sí la*

cantidad de valores que pueden asumir las tres variables (o sea $2 \times 2 \times 2$)⁴⁴. Puede apreciarse cómo crece el número de grupos a medida que incluimos más variables. Cuando al diseño de 2×2 le agregamos una tercera, aumenta de 4 a 8 y si lo hacemos con una cuarta, es de 16 ($2 \times 2 \times 2 \times 2$). Esto muestra una de las principales dificultades que tienen para ser implementados: la enorme cantidad de grupos que debemos formar, cualquiera sea el tipo al que pertenecen (cuasiexperimental, no experimental o riguroso).

Diagrama 5.5. Diseño factorial sin distribución al azar, $2 \times 2 \times 2$, con pretest y posttest

t_1	t_2	t_3	t_4
$N G_1$	O_{11}	$V_0 W_0 X_0$	O_{21}
$N G_2$	O_{12}	$V_0 W_0 X_1$	O_{22}
$N G_3$	O_{13}	$V_0 W_1 X_0$	O_{23}
$N G_4$	O_{14}	$V_0 W_1 X_1$	O_{24}
$N G_5$	O_{15}	$V_1 W_0 X_0$	O_{25}
$N G_6$	O_{16}	$V_1 W_0 X_1$	O_{26}
$N G_7$	O_{17}	$V_1 W_1 X_0$	O_{27}
$N G_8$	O_{18}	$V_1 W_1 X_1$	O_{28}

Validez interna y tipo de diseño

El Cuadro 5.2 muestra los factores de validez interna que incluyeron Campbell y Stanley (2005) y repiten Shadish et al. (2002), aunque uno de ellos que no fue mencionando hasta ahora, se observa sólo en el último trabajo: *la ambigüedad de la precedencia temporal* (ambiguous temporal precedence). Sobre ella que sostienen que “la falta de claridad acerca de qué variable se produjo primero puede producir confusión sobre cuál variable es la causa y cuál es el efecto” (: 55). Este factor es

⁴⁴ Este diseño, como los restantes presentados de tipo cuasiexperimental, también podría realizarse en forma rigurosa, si los sujetos se distribuyeran al azar en los distintos grupos.

particularmente importante en los diseños no experimentales en los que en muchas oportunidades no se puede establecer qué ocurrió en primer lugar. Sólo el conocimiento del área de investigación pueden en algunos casos salvarnos pero en otros esto resultará imposible.

Cuadro 5.2. Definición de los factores de validez interna.

Ambigüedad de la prece- dencia temporal	No existe seguridad de que la variable indepen- diente precede a la dependiente.
Historia	Sucesos que se desarrollan en forma paralela al experimento
Maduración	Procesos de cambio interno de los sujetos durante el experimento
Administración de test (o pruebas)	Procesos de aprendizaje o modificaciones que produce el hecho de responder a pruebas
Instrumentación	Cambios en la forma en que los instrumentos de recolección de datos miden la variable dependiente o de interpretar lo observado
Distribución	Falta de asignación al azar de los sujetos al grupo control y experimental
Mortalidad experimental	Abandono del experimento o pérdida de datos en forma diferente en los grupos de control y experimental, o luego del pretest de los casos con valores menos favorables en relación a la variable dependiente
Regresión	Tendencia de los puntajes extremos de regresar a la media en una segunda medición
Interacción	Interacción entre distribución y otros factores de validez

Otro punto importante se relaciona a una tabla que incluyeron Campbell y Stanley (2005) en la que muestran los factores de validez (ubicados en las columnas) que controlan cada diseño (en las filas). El signo “+” en una celda indica que ese factor es controlado. Pero algo sumamente sugerente es que en el trabajo de Shadish et al. (2002) estas tablas no están presentes. Veamos algunas ideas que marcan una diferencia de concepción entre ambos trabajos.

Shadish et al. (2002) sostienen que “la asignación aleatoria simplemente reduce la probabilidad de que esas amenazas [de validez interna] sean confundidas con el tratamiento”. Así, resulta claro que, si bien esta forma de control (la asignación al azar) es un poderoso instrumento para reducir las amenazas, no las elimina por completo. Incluso estos autores llegan a sostener que hay continuidad entre los diseños experimentales aleatorizados y los no experimentales, pasando por los cuasiexperimentales y que todos pueden ser plausibles. Además, afirman que la validez depende de las características del estudio: por ejemplo, se hizo un estudio con dos grupos formados al azar, pero en uno de los grupos notamos que abandonaron el 75% mientras que en el otro lo hizo sólo el 25%⁴⁵, ¿podemos confiar en que la mortalidad no es una amenaza? Además, la mortalidad puede darse por el abandono diferencial en los grupos, si quienes dejan el experimento en el grupo de control son los más motivados y en el experimental los de menor motivación, esto provocará una diferencia no debida a la variable independiente cuyo efecto se confundirá con el de ella⁴⁶. Es poco probable que se dé un abandono tan diferente en un estudio aleatorizado, pero el ejemplo sirve para mostrar que debemos analizar las amenazas en relación a las situaciones concretas que se presentan. Shadish et al (2002) también señalan que la mortalidad puede verse como un sesgo de asignación que se produce luego de comenzar el experimento, y no es controlada por la asignación al azar, esto último marca una diferencia con el trabajo de Campbell y Stanley (2005) cuyo original en inglés se publicó en 1966 (aquí utilizo la novena reimpresión en castellano).

Otra idea interesante en la que insisten, es que las amenazas de validez son más o menos plausibles según los resultados que se observen. Ponen el ejemplo hipotético ya mencionado de un estudio con el que llamo *Diseño de un grupo con posttest únicamente*. Supongamos que probamos una propuesta para que los niños lean más en vez de jugar con la computadora (por ejemplo, implementada en la escuela). Puesto que conocemos que la mayoría ellos se interesa poco por la lectura, y mucho por el ordenador, si luego de aplicar el tratamiento encontramos un cambio notable en los niños, podemos atribuírselo al tratamiento, pues es poco probable que éste se dé por otros factores como la historia o la maduración, etc. Así, si el efecto es contundente por un lado y poco probable de observarse en la muestra estudiada, por otro lado,

45 Menciono un caso muy poco probable para ilustrar el concepto.

46 Puede ser otra variable distinta de la motivación la que produce la diferencia.

entonces podemos atribuírselo al tratamiento con menos posibilidades de equivocarnos.

Esto sugiere que los factores de validez no deben interpretarse relacionados a la forma de cada diseño como la Tabla del trabajo de Campbell y Stanley (2005) parece sugerir (aunque en realidad ellos nunca quisieron sostener esta idea⁴⁷).

Volvamos al ejemplo mencionado más arriba (la comparación entre dos métodos de enseñanza de la lectoescritura), que consistía en enseñarle al primer grado *A* del turno mañana con el primer método (*M1*), mientras que al primer grado *B* del turno tarde con *M2* y, luego de transcurridos 6 meses de enseñanza, tomarle una prueba de lectoescritura a ambos grupos. Supongamos que en lugar de hacerlo de esta forma, decidimos conformar dos grupos al azar, pero en el grupo experimental enseña una docente muy motivada, que establece una buena relación con los alumnos y es creativa a la hora de implementar la enseñanza. En el grupo de control, la docente tiene características opuestas (esto podría suceder aún si la determinación de cuál es la docente asignada a cada grupo se realiza al azar). Así, este factor podría provocar una diferencia entre los grupos que nada tiene que ver con *X*. Surge, entonces, una pregunta inevitable: ¿es la lista que nos presentan Shadish et al. (2002) exhaustiva? Ellos mismos sostienen que no, los factores de validez señalados son los que más probablemente actuarán en determinados diseños e insisten en que se deben entender como los *cuidados o precauciones que deben tomar los investigadores, antes que pensar que las inferencias realizadas a partir de la forma del diseño permiten de un modo determinista interpretar los resultados de un experimento*.

En otros términos, los estudios rigurosos no nos garantizan en forma absoluta el control de la validez interna. Así, el análisis de estos autores *nos debe servir para reflexionar sobre las posibles dificultades de nuestros*

47 Discuto estas cuestiones en Echevarría (2011b y c). En ellos argumenté que de Campbell y Stanley (2005) pueden hacerse dos interpretaciones, una fuerte y otra débil. Según la primera “hablamos fundamentalmente de factores de validez cuya falta de control hace que los diseños la tengan en mayor o menor grado. Pero una lectura atenta de esta obra [Campbell y Stanley, 2005] pone de manifiesto que ellos no adoptan una posición tan rígida, pues en varios pasajes parece predominar una concepción que denominaré *débil*. Aquí consideramos que más que de factores, estamos ante amenazas de validez (...). O sea que la lista que se discute (cualquiera sea ésta) nos indica posibles amenazas que podrían invalidar las inferencias” (Echevarría, 2011b: 183). Creo que en el trabajo de Shadish et al. (2002) se nota con mayor nitidez la interpretación débil, lo que explica la falta de la tabla de los factores de validez presente en Campbell y Stanley (2005).

experimentos, antes que como normas o cánones interpretativos que necesariamente se darán ante tal o cual diseño.

Otro punto también espinoso queda por tratar: los diseños rigurosos y cuasiexperimentales son fuertes en validez interna, pero logran este cometido por caminos muy diferentes. En el primer caso, por la aleatorización que permite la neutralización de los factores ajenos a la variable independiente, es decir, se trata de evitar que estos factores influyan en Y (variable dependiente) o lo hagan de un modo diferente en el grupo experimental y de control. En cambio, los cuasiexperimentales solamente estiman (miden) si estos factores están actuando sobre Y , lo cual resulta radicalmente diferente (recuérdese lo que se dijo en el ejemplo que incluí al introducir el tema de la validez interna). Esto significa que el análisis de la forma de los diseños nos muestra cómo podemos mejorar nuestros experimentos de un modo sencillo con modificaciones que, generalmente, no implican un gran gasto de tiempo o recursos adicionales.

Finalmente, veamos otro aspecto que habíamos mencionado más arriba y que quedó pendiente. Como ya dije al hablar de los diseños rigurosos, según algunos autores, en la investigación experimental se debe cumplir el requisito del doble ciego. ¿Qué pasa entonces si esto no se da?, ¿el diseño carece de validez? Creo que no, pero es otro factor a tener en cuenta, tomando la mayor cantidad de precauciones posibles. Por ejemplo, si está llevando a cabo su investigación en el aula, asignando al grupo con exposición a X_i a aquellos docentes que no creen en las innovaciones propuestas o que están menos predispuestos a implementarlas, podría repetirlo con otros alumnos invirtiendo los educandos asignados a cada grupo.

Con todo, digamos que los diseños rigurosos, en términos generales, son los más fuertes en validez interna, teniendo el máximo nivel de control, tanto los de una variable independiente, como los factoriales. Por esta razón, es deseable su uso, al menos cuando el investigador ya ha clarificado sus hipótesis y ha probado sus instrumentos de recolección de datos. Si su teoría se vio corroborada por los hechos, en una etapa posterior podría intentar una nueva contrastación con estudios cuasiexperimentales, buscando en este caso una mayor aproximación a la validez externa.

De todos modos, no debe olvidarse que cuando no es posible la aleatorización, siempre pueden mejorarse los diseños agregando mediciones o grupos de comparación. Comencé con los estudios cuasiexperimentales de un grupo, y luego presenté aquellos que tienen dos o más de ellos, pero en todos los casos, expuse primero los más básicos, los que tienen menos mediciones y por lo tanto mayores amenazas de validez interna para luego mostrar cómo se pueden agregar observaciones o grupos de comparación (o ambas cosas) y de este modo aumentar las posibilidades de control de la validez⁴⁸.

Y por último, recuerdo una idea de Shadish et al. (2002) que por su importancia la reitero: cuando el efecto observado es contundente, podemos hacer algunas concesiones en cuanto a la rigurosidad del diseño, en cambio cuando éste no es demasiado claro, su forma nos puede dar alguna garantía en la interpretación de los resultados. De todos modos, un experimento e incluso unos pocos, no son suficientes para aceptar una hipótesis dada.

48 Otra diferencia entre los trabajos de Shadish et al. (2002) y Campbell y Stanley (2005) es la desaparición de los diseños preexperimentales, como dije en la primera edición de este libro la diferenciación entre ambos es difícil de establecer. También agregan un factor de validez: precesencia temporal ambigua (ambiguos temporal precedente). Por otro lado, no todos los autores que tratan el tema dan la misma lista aunque en general parten de Campbell y Stanley (2005).

Capítulo 6

Investigación descriptiva

Los *estudios descriptivos* son particularmente útiles cuando un investigador se inicia en un tema nuevo. En este caso, puede comenzar a recabar datos vinculados al problema o tema que recortó, sistematizarlos y exponerlos, sin pretender establecer relaciones de causalidad entre variables.

Aquí trato los *estudios descriptivos*⁴⁹ que usan datos cuantitativos y muestran con qué frecuencias aparece cada categoría para las distintas variables, qué valores asumen ciertos parámetros que caracterizan a una población o cómo evolucionan estos valores en el tiempo. Esto último puede posibilitar realizar predicciones o conducirnos a descubrir relaciones entre variables, lo que marca el límite con la investigación no experimental, pues al establecer estas relaciones ya estamos pensando en términos de la asociación que puede haber entre ellas.

Los diseños descriptivos resultan de utilidad para conocer quiénes, cuántos, cuándo, cómo, qué cambios se producen con el tiempo o qué diferencias hay entre ciertos grupos en relación con determinadas variables, y están siempre referidos a una situación acotada y a un tiempo

49 Para algunos, la investigación cualitativa es descriptiva (por ejemplo, Bogdan y Biklen, 1992); dejando de lado este tópico, en este capítulo trato la investigación descriptiva de carácter cuantitativo.

determinado. Pueden responder, por ejemplo, a preguntas como las siguientes: ¿qué libros leen los adolescentes?, ¿quiénes estudian carreras humanísticas?, ¿cuántos alumnos tienen dificultades de aprendizaje?, ¿cuándo se produce la mayor deserción en la universidad?, ¿cómo evolucionan los sujetos sometidos a un tratamiento psicopedagógico determinado?, ¿qué diferencia hay entre los docentes de mayor antigüedad y los de menor experiencia en la forma de enseñanza? Estas preguntas se realizarán en relación con una población y en un tiempo estipulado⁵⁰.

Por otro lado, las diferencias entre un estudio descriptivo y uno explicativo también se aprecian en el hecho de que en el primero carece de sentido hablar de variables independientes y dependientes, pues no se establecen relaciones de causalidad entre ellas. Puesto que no se contrastan hipótesis causales, propongo clasificarlas en *descriptoras*, que son aquellas que se usan para caracterizar a los sujetos que se intenta describir, y de *comparación*, que nos permiten diferenciar entre grupos. Existe una analogía entre algunos estudios descriptivos y los de contrastación de hipótesis no experimentales. Las variables de comparación en los primeros ocupan el mismo lugar que las independientes en los que se ponen a prueba hipótesis, y recíprocamente, las variables descriptoras y dependientes cumplen roles análogos en unos y otros⁵¹.

Los diseños transversales

El diseño descriptivo más elemental es el de un grupo y una sola medición (Diagrama 6.1). Se realiza un corte *transversal* en el tiempo y

50 Lo más común es utilizar frecuencias, medidas de tendencia central (como la media aritmética, la mediana o la moda) y medidas de dispersión (como por ejemplo, la desviación estándar). Pero no tenemos porqué limitarnos a estos estadísticos.

51 Cohen y Manion (1990) conciben la investigación descriptiva como una clase amplia que incluye, en primer lugar, los estudios longitudinales, transversales y de tendencia o predicción. Pero también ven como tales a los estudios correlacionales (aquí llamados no experimentales), el estudio de casos, la investigación histórica, la encuesta y la investigación ex post facto. Esto es así ya que ellos consideran dos grandes categorías: por un lado, la investigación experimental, y por otro lado, la descriptiva. Arnal et al. (1992), al hablar de las metodologías de corte cuantitativas diferencian entre experimental, cuasiexperimental y no experimental y la última contiene a la correlacional y descriptiva. Sin embargo, considero que a ésta es preferible ubicarla en una categoría aparte, porque tiene objetivos que la diferencian claramente, tanto del estudio de casos, como de la investigación que aquí llamo no experimental. Las denominaciones y la concepción de los diseños descriptivos que incluyo están tomadas de Manheim (1982), aunque formalizados con la simbología usada aquí, pues, como ya dije, me propuse representarlos a todos con las mismas convenciones.

se mide la variable descriptora. También podría suceder que se consideren más variables, pero igualmente hablamos de una observación, en el sentido de que se indagan en una única oportunidad. El objetivo presentado más arriba responde a este diseño (conocer el rendimiento de los alumnos de Ciencias Económicas que cursaron las materias Lógica y metodología de las ciencias en el primer cuatrimestre del año 1999).

Puede verse que en este diagrama se incluye una sola variable, pues he hecho una descripción no comparativa. Además, se aprecia que los estudios descriptivos no analizan relaciones de causalidad entre variables y pueden presentarse los datos con una sola variable por vez. Pero esto no necesariamente es así, por ejemplo, si queremos hacer una descripción de la nota según el sexo de los alumnos, podemos construir el gráfico que se presenta más adelante al tratar el diseño transversal de dos grupos.

El diseño de un grupo y una sola medición sólo nos permite conocer las características de una población en un momento dado. En cambio, si le agregamos uno o más grupos, además de esto podremos realizar una comparación entre ellos, como se aprecia en los Diagramas 6.2 a 6.4⁵².

Diagrama 6.1. Diseño transversal de un grupo

$$\begin{array}{cc} t_1 & t_2 \\ N G_{x1} & O \end{array}$$

Diagrama 6.2. Diseño transversal de dos grupos

$$\begin{array}{cc} t_1 & t_2 \\ N G_{x1} & O_1 \\ N G_{x2} & O_2 \end{array}$$

Por ejemplo, si además de conocer el rendimiento de los alumnos en el primer parcial, quisiéramos comparar entre mujeres y varones tene-

⁵² A veces suelen llamarse a estos diseños seccionales, reservando el término transversal para aquellos que estudian grupos de edades distintas con la finalidad de conocer su evolución. Es decir, permiten lograr lo mismo que los longitudinales, aunque en un período de tiempo breve.

mos un diseño transversal de dos grupos. En este caso, debemos incluir en el diagrama las dos variables, que serviría para tres objetivos: conocer el rendimiento en el primer parcial de lógica de los alumnos de sexo femenino, de los de sexo masculino y comparar entre ambos. Hemos usado un diseño *transversal* (los datos fueron recolectados en una única oportunidad y en dos grupos) y *comparativo*, pues nos permite comparar entre ellos.

A su vez, podemos implementar diseños transversales tres o cualquier número de grupos, como se ve en los Diagramas 6.3 y 6.4.

Diagrama 6.3. Diseño transversal de tres grupos

t_1	t_2
NG_{x1}	O_1
NG_{x2}	O_2
NG_{x3}	O_3

Diagrama 6.4. Diseño transversal de n grupos

t_1	t_2
NG_{x1}	O_1
NG_{x2}	O_2
.	.
.	.
.	.
NG_{xn}	O_n

Diseños longitudinales

Los *diseños longitudinales* se caracterizan por realizar un seguimiento de uno o más grupos, concretando varias mediciones de las mismas variables en distintos momentos, lo que facilita conocer si se van mo-

dificando o no a través del tiempo. Si estos diseños tienen más de un grupo, a su vez, además de apreciar su evolución, se pueden hacer comparaciones como en los transversales (Diagramas 6.5 a 6.8).

Con este tipo de estudios, a veces es imposible ubicar a los mismos sujetos, por lo que Hernández Sampieri et al. (1991) diferencian entre diseños de panel, de evolución de grupo y de cohorte.

En el primer caso, se considera exactamente a las mismas personas en todas las observaciones, mientras que en los restantes, las muestras que se toman son equivalentes, pero no necesariamente coinciden todos los sujetos seleccionados (puede ocurrir que algunos se incluyan en más de una medición, otros en una sola y algunos en todas). Pero si todos los individuos cambian, estamos en presencia de un tipo de diseño que veremos en el siguiente apartado (y aquí, siguiendo a Manheim, 1982, los llamo de *etapas emparejadas*).

Diagrama 6.5. Diseño longitudinal de un grupo y dos observaciones

$$\begin{array}{ccc} t_1 & t_2 & t_3 \\ N G_{x1} & O_1 & O_2 \end{array}$$

Diagrama 6.6. Diseño longitudinal de un grupo y n observaciones

$$\begin{array}{ccccccc} t_1 & t_2 & t_3 & \cdot & \cdot & \cdot & t_{n+1} \\ N G_{x1} & O_1 & O_2 & \cdot & \cdot & \cdot & O_n \end{array}$$

Diagrama 6.7. Diseño longitudinal de dos grupos y n observaciones

$$\begin{array}{ccccccc} t_1 & t_2 & t_3 & \cdot & \cdot & \cdot & t_{n+1} \\ N G_{x1} & O_{11} & O_{21} & \cdot & \cdot & \cdot & O_{n1} \\ N G_{x2} & O_{12} & O_{22} & \cdot & \cdot & \cdot & O_{n2} \end{array}$$

Diagrama 6.8. Diseño longitudinal de n grupos y n observaciones

t_1	t_2	t_3	.	.	.	t_{n+1}
NG_{x1}	O_{11}	O_{21}	.	.	.	O_{n1}
NG_{x2}	O_{12}	O_{22}	.	.	.	O_{n2}
.	.	.	.	.	.	.
.	.	.	.	.	.	.
.	.	.	.	.	.	.
NG_{xm}	O_{1m}	O_{2m}	.	.	.	O_{nm}

Diseños de etapas emparejadas

Cuando no se puede seguir el mismo grupo, pero se lo compara con otro considerado equivalente, tenemos un *diseño de etapas emparejadas*. Un ejemplo podría ser el siguiente: se desea conocer si el porcentaje de deserción universitaria al finalizar el primer año de estudio, en la Argentina, se modificó o permaneció estable luego de la crisis que sufrió el país desde la finalización del año 2001. Para tal fin, podemos tomar una muestra de ingresantes del año 2000, otra del 2001, del 2002 y del 2003 (todas seleccionadas de la misma forma) y comparar los porcentajes de deserción entre ellas. En este caso, no tiene sentido observar la variable en el mismo grupo, pues nos alejaría de nuestro objetivo, el cual es conocer la deserción al finalizar el primer año de cursado, y la cohorte 2000 ya superó ese tiempo en el 2001, con lo que deja de tener la característica que nos interesa.

Con estos diseños, el número mínimo de grupos es dos, pero el máximo es indefinido. Notemos que, si bien todos son diferentes, tienen una característica en común (son equivalentes desde este punto de vista), señalada en los diagramas con x_j . A modo de ejemplo, presento el de dos, tres y n grupos (Diagramas 6.9 a 6.11).

Diagrama 6.9. Diseño de etapas emparejadas de dos grupos

t_1	t_2	t_3
NG_{x1}	O_1	
NG_{x1}		O_2

Diagrama 6.10. Diseño de etapas emparejadas de tres grupos

t_1	t_2	t_3	t_4
NG_{x1}	O_1		
NG_{x1}		O_2	
NG_{x1}			O_3

Diagrama 6.11. Diseño de etapas emparejadas de n grupos y n observaciones

t_1	t_2	t_3	.	.	.	$tn+1$
NG_{x1}	O_1	.	.	.	.	.
NG_{x1}		O_2	.	.	.	.
.	.	.	.	.	.	.
.	.	.	.	.	.	.
.	.	.	.	.	.	.
NG_{xm}	.	.	.	.	.	O_n

La investigación descriptiva y el problema de la validez

Al igual que en los estudios factoriales, el problema de la validez no ha sido tratado para los diseños descriptivos con el mismo grado de sistematicidad con que se lo hizo para la investigación experimental.

Los *diseños longitudinales* pueden tener dificultades con la *mortalidad*, debido a que suele ser difícil hallar a las mismas personas en todas las oportunidades en que se decidió recabar información y tendrá más importancia cuando mayor sea el número de sujetos que se modifican entre una y otra observación. De todos modos, el investigador puede estimar su efecto, pues conoce cuántos son los casos que variaron entre las distintas mediciones.

Otro factor difícil de controlar es la *instrumentación*, pues suelen cambiarse los entrevistadores o la forma de interpretar las respuestas obtenidas. La *administración de test* influye en todo estudio en que se hagan

mediciones sucesivas a un mismo grupo (recordemos que se relaciona con el proceso de aprendizaje o las modificaciones que desencadena la sola aplicación de una prueba), por eso, afecta la validez de los diseños longitudinales. A su vez, la *regresión* actuará si el grupo seleccionado tiene un nivel bajo o alto, con lo cual en las mediciones siguientes podríamos registrar valores más altos o bajos respectivamente, aunque esto es improbable si la muestra se extrajo al azar.

La *historia* (definida como todos los cambios que se producen mientras dura la investigación) difícilmente pueda intervenir. Por ejemplo, si queremos realizar un estudio longitudinal en un colegio y elegimos justamente un año en que se están implementando reformas edilicias de importancia (con las tensiones que puede originar), podríamos hacernos una idea equivocada acerca de cómo es la realidad del mismo, o de la evolución que tuvieron ciertas variables en un lapso de tiempo determinado. Sin embargo, puesto que cuando describimos sólo estamos mostrando la realidad particular de un grupo de sujetos en un momento determinado, considero que no afecta la validez de una manera crucial. En todo caso, es conveniente registrar con la mayor fidelidad posible los sucesos que se fueron dando en forma paralela a la investigación, para dejar establecido que estos valores de las variables descriptoras se observaron bajo esas condiciones particulares.

Respecto de los *estudios transversales*, es poco probable que la *historia* actúe en ellos, pues se recolectan los datos en una sola oportunidad. Además, debido a que no realizamos mediciones sucesivas de las mismas variables, no existe la posibilidad de perder sujetos en el tiempo (*mortalidad*), de que se produzca aprendizaje por la implementación de los instrumentos de medición en forma repetida (*administración de test o pruebas*), ni modificaciones en su aplicación o en la interpretación de lo obtenido (*instrumentación*). Como contrapartida, ofrecen menor cantidad de información que los longitudinales, siendo imposible analizar con ellos la evolución de ciertas variables, como ocurre con los últimos.

Los *diseños de etapas emparejadas*, a su vez, tienen algunas dificultades de los longitudinales, ya que en realidad son una variación de ellos, aunque controlan la administración de test, pues en ningún caso se observa en forma sucesiva a los mismos sujetos como en los últimos.

Ejercicios

1. Del Apéndice B, selecciona él o los estudios descriptivos y realiza el diagrama del diseño.
2. Pensá en una investigación descriptiva qué podrías realizar en tu Trabajo Final (o en alguna otra exigencia académica) y describila en media página
3. Realizá el diagrama de su diseño.

Consideraciones finales

¿Qué diseño elegir?

En la elección del diseño de investigación existen varias consideraciones relevantes. Uno de ellos se relaciona a la necesidad de modificar el normal funcionamiento de la institución en que se podría llevar a cabo la investigación. En general los diseños cualitativos⁵³, los descriptivos y los no experimentales, suelen ser más fáciles de implementar que los experimentales. Los cualitativos por un lado tienen la ventaja de que no hace falta modificar el ambiente como en los experimentales, pero los procesos de recolección y análisis de datos son sumamente trabajosos. En efecto, si se trata de un estudio de casos con una permanencia prolongada en el escenario, puede resultar extremadamente difícil de llevar a la práctica, sobre todo por el tiempo que insumen las observaciones y el análisis de los datos resultantes. Los rigurosos en general son difíciles de implementar por la necesidad de romper los grupos naturales y formarlos al azar, y todos los factoriales por la cantidad de grupos que requieren.

⁵³ Recuérdese que los diseños cualitativos no se tratan aquí, pues como ya dije, le dediqué un libro a ellos.

El aspecto más importante para la elección del diseño está determinado por los objetivos que persigue el investigador. El Cuadro 7.1 nos muestra aquéllos a los que pueden contribuir los distintos modelos. Sin embargo, se debe tener presente que cada una de las afirmaciones que se incluyen en las celdas indican una orientación general que en modo alguno representan una valoración adecuada para todos los casos concretos que podemos considerar. Se incluyeron aquellos a los que más frecuentemente recurren los investigadores con cada modelo, pero aquí también debemos tener en cuenta que no todos coinciden en este punto. Por ejemplo, para algunos investigadores con estudios cualitativos pueden ponerse a prueba hipótesis causales, mientras que para otros, esto implica desvirtuarlos por completo.

Cuadro 7.1. Finalidad con que suelen utilizarse y manipulación que se da en los diseños de investigación

Diseño	Finalidad	Manipulación
Cualitativos	6,7	N
Transversal	1,2,3,7	N
Longitudinal	1,2,3,7	N
Etapas emparejadas	1,2,3,7	N
No experimentales	2,3,4,7	N
Cuasiexperimentales	4	S
Rigurosos	4	S
Factorial cuasiexp.	4,5	S
Factorial riguroso	4,5	S
N=1	4,7	S
Referencias: 1. Establecer cómo es la realidad; 2. Descubrir relaciones entre variables; 3. Hacer predicciones; 4. Contrastar hipótesis causales; 5. Analizar interacción entre variables; 6. Interpretar casos que tienen una cultura muy diferente a la del investigador; 7. Construir un diseño u otros aspectos para otra investigación. Lo que se sostiene en el cuadro no es taxativo; indica lo más frecuente para cada tipo de estudio.		

Respecto de la validez externa, en general los diseños que no suponen manipulación se consideran fuertes desde este punto de vista, y débiles los estrictamente experimentales (por la reactividad que normalmente presentan), pero esto podría no ser así para algunos estudios específicos en situaciones determinadas, e iguales consideraciones podrían hacerse con la validez interna.

Por otro lado, hay otras cuestiones de importancia a la hora de tomar una decisión. Una de ellas es el costo, y correlativamente, los recursos disponibles, que no sólo tienen que ver con cuestiones puramente económicas, sino también con otros aspectos. Por ejemplo, para hacer una investigación cualitativa, se necesitan observadores que tengan experiencia en ella y un cierto conocimiento del marco conceptual que se usará, cosa que no es relevante en una encuesta, en la cual quienes recolectan la información hasta pueden ignorar los objetivos con que se lleva a cabo.

En relación con las posibilidades de aplicación de cada tipo de estudio, si el investigador se desempeña en la clínica, o tiene acceso a casos individuales (por ejemplo un profesor que da clases particulares), puede apelar al diseño de sujeto único o estudio de casos. Si quiere poner a prueba una teoría, y en un colegio le permiten formar grupos a voluntad, puede realizar una investigación experimental rigurosa o factorial (esto último dependerá del número de grupos que se puedan concretar); pero si no es factible organizarlos como él lo desea, cuenta con los diseños cuasiexperimentales, que pueden ser de uno, dos, o más grupos, según lo que la institución le posibilite y sus objetivos le señalen.

También se podría llevar a cabo una investigación cualitativa como forma de preparar otra de carácter cuantitativo. Una de las formas en que suele hacerse consiste en realizar un sondeo exploratorio con el propósito de precisar el problema, construir los instrumentos de indagación, seleccionar las técnicas de análisis de los datos y formular las hipótesis de trabajo, para luego realizar otro tipo de estudio. Así, si el investigador no tiene en claro estos aspectos esenciales de su proyecto, puede en una primera fase recolectar datos con instrumentos inestructurados, y luego codificar los registros narrativos, desarrollando categorías de bajo nivel. Si estas categorías son muy numerosas, pueden unirse algunas formando clases. Finalmente, el cálculo de frecuencias, la aplicación de algún procedimiento de análisis multivariado exploratorio y la formulación de hipótesis, pueden coronar el trabajo, con lo

cual contará con algunas conclusiones, además de preparar el terreno para llevar a cabo en el futuro algo más estructurado. También los datos cualitativos pueden servir de complemento a los cuantitativos, usando extractos de los registros narrativos con la finalidad de ejemplificar algunas de las categorías de análisis establecidas⁵⁴.

¿Otros tipos de investigación?

Si el lector revisa la bibliografía sobre los tipos de investigación posibles, encontrará otros que hasta ahora no han sido mencionados o lo han sido sólo tangencialmente, como el seguimiento de cohorte, la encuesta, los estudios de diagnóstico, la investigación acción, la clínica, la evaluativa, y otros más. *Pero todos pueden seguirse de acuerdo a algunos de los modelos en forma pura o combinando dos o más de ellos*. Por ejemplo, la encuesta a veces responde a una lógica descriptiva, pero en otras es explicativa no experimental. A su vez, estas dos alternativas pueden ser una parte de un estudio de diagnóstico, que además podría albergar en su interior uno experimental o incluso cualitativo.

Al igual que la encuesta, el seguimiento de cohorte por lo general es descriptivo. La investigación acción puede ser de cualquiera de los tipos ya mencionados, aunque se la suele asociar a los cualitativos y se le agrega alguna intervención, que es lo que la caracteriza; en la clínica se han realizado estudios documentales, experimentales, encuestas, cualitativos y de investigación acción (Miller y Crabtree, 1994). Todos ellos se caracterizan porque tienen una finalidad que por lo general es más amplia que la de los presentados en la clasificación del Cuadro 7.1. Por ejemplo, la investigación evaluativa se realiza para describir la situación que se quiere evaluar, pero también es conveniente explicarla, valorarla y realizar recomendaciones para el futuro. Por esta razón, podría incluir un estudio descriptivo para el primer objetivo, uno experimental o no experimental para los otros.

⁵⁴ Esto no significa que la investigación cualitativa siempre sea exploratoria, pues en algunos casos puede representar el interés principal del investigador. Incluso una investigación cuantitativa podría ser preparatoria de ella.

Lógica en uso y lógica reconstruida

La finalidad de este trabajo no es dar una visión detallada de cada tipo de estudio, sino un panorama global que permita una aproximación introductoria al tema y, al mismo tiempo, plantear un primer análisis de la validez en los casos en los que no se lo ha hecho sistemáticamente en la bibliografía relevada, aunque limitada a los típicos diseños de la tradición cuantitativa. Haré algunas consideraciones sobre las circunstancias en que puede resultar adecuado usar una única opción, como así también la combinación de propuestas metodológicas provenientes de alternativas diferentes.

El investigador no debe usar estos diseños rígidamente sino, al contrario, de un modo flexible, y siempre teniendo en cuenta el grado de avance de su proyecto y los problemas que trata de resolver en cada etapa. Así, es posible que en un primer momento -cuando está clarificando el problema, construyendo el marco teórico y los instrumentos de recolección y de análisis de los datos- resulte adecuado emprender un estudio descriptivo, para pasar luego a uno no experimental y finalmente a uno experimental. Pero esta secuencia, si bien tiene su sentido, no necesariamente debe ser seguida por todos. Podría resultarle conveniente a alguien que tiene por principal finalidad desarrollar una teoría causal sobre un determinado fenómeno, pero poco provechosa para otros que pretendan comprender la forma de vida de una comunidad determinada. En este caso, podría ser conveniente comenzar con una encuesta descriptiva, que permita conocer de un modo general a la población y, a partir de esta visión global, seleccionar casos para avanzar hacia una investigación cualitativa. O quizás comenzar con la última y luego con la encuesta indagar hasta qué punto las interpretaciones realizadas pueden generalizarse de la muestra con que se trabajó al universo que representa.

Esto significa que me inclino fuertemente en favor de una integración de las distintas alternativas metodológicas. Pero también por una plena libertad del investigador para elegir el método que le resulte más adecuado en cada momento de su proyecto, lo que nos lleva al debate sobre la lógica que subyace en la tarea de investigar. Podemos distinguir, por un lado, la *lógica en uso*, que es aquella que se pone en práctica, a veces sin conciencia de su existencia y, por otro lado, la *lógica reconstruida*, que surge de una reflexión sobre la primera y representa una

idealización de lo que hace el investigador (Gibaja, 1988). Lo que aquí presenté corresponde a la última que, como ya sostuve, no debe ser prescriptiva, o no al menos de un modo taxativo. Gibaja (ob. cit.) nos señala el peligro de magnificar estas idealizaciones: puede llevarnos a posiciones intransigentes, que surgen de la adhesión dogmática a algún tipo de estudio.

La lógica reconstruida tiene la importante finalidad de funcionar como un instrumento heurístico, que le muestra al investigador algunas alternativas, como así también los puntos fuertes y débiles que surgen de analizar la experiencia acumulada en cada uno de estos modelos. Si cuenta con algunas de estas posibilidades, dispone de una serie de instrumentos probados y analizados para desarrollar su trabajo. Esto es interesante, pero no debemos dejar de tener presente que es conveniente analizar permanentemente la pertinencia de cada metodología al problema que se enfrenta y, si el investigador lo cree conveniente, salirse del camino trazado por el metodólogo y seguir el que le *indica su propia lógica*. Después se puede discutir si lo que hizo tiene o no fundamentos, *enriqueciendo la lógica reconstruida*, cuya utilización flexible puede *llevar a una nueva lógica en uso*, y así sucesivamente.

Por otro lado, la lógica reconstruida es siempre incompleta, a veces simplemente equivocada, otras contradictoria. Quizás un ejemplo de lo segundo podemos hallarlo en las críticas que se le hicieron a Piaget, de quien entre otras muchas cosas se dijo que no buscaba explicaciones causales en sus investigaciones, cuando en realidad él nunca se lo propuso (Vuyk, 1984). Tal vez un ejemplo de lo último (lógicas reconstruidas contradictorias) podemos encontrarlo en el uso de pruebas estadísticas en experimentos de sujeto único. Mientras que algunos autores piensan que esto es conveniente, otros no lo recomiendan. ¿Qué hace entonces el investigador? Lo más sensato es analizar las razones por las que unos las rechazan y otros las aprueban y ver en qué medida estos argumentos le sirven para realizar una crítica a su propio trabajo y de ese modo mejorarlo, antes que seguir rígidamente un camino trazado por alguna de las lógicas reconstruidas.

Esto no significa caer en una posición anárquica en la que todo vale, pues, una vez adoptada una alternativa, debemos ser coherentes con los fundamentos que ella conlleva, como así también con las estrategias y técnicas que supone y, si adoptamos una perspectiva ecléctica o una posibilidad no explorada aún, no debemos dejar de reconocerla y

plantearnos permanentemente tanto sus fundamentos como los puntos fuertes y débiles que pudiera presentar.

Quizás el lector advierta en estas palabras que veo a la metodología como una disciplina orientada a ayudar al investigador a tomar las imprescindibles decisiones para avanzar en su trabajo (aunque no tenga ni se crea con el deber de indicarle un único modo de hacerlo), o más aún, como una fuente de inspiración que le sugiere posibilidades en las que tal vez no pensó, antes que un juez implacable cuyas prescripciones deben seguirse ciegamente, sin la más mínima posibilidad de desviación o de crítica.

Apéndice A

Tabla A1. Elementos que podemos hallar en los antecedentes y que son relevantes para el planteo del problema o de las preguntas de investigación

Autor, Título	Supues- tos	Concep- tos	Hipóte- sis	Datos	Proble- ma social	Expe- riencias, reco- menda- ciones para la acción

Apéndice B. Ejemplos

González et. al. (1998) realizaron un estudio de las conductas de alta y baja sociabilidad en forma simultánea en Buenos Aires y Valencia. El objetivo central fue “conocer las posibles diferencias existentes cuando se toman como variables moduladoras el sexo y el lugar de residencia, dejando como constantes las variables edad cronológica (13 a 18 años) y nivel educativo” (p. 132).

En las dos ciudades utilizaron un muestreo intencional y como instrumento de recolección de datos usaron dos cuestionarios cerrados. El primero, denominado Cuestionario de Conducta Prosocial fue elaborado por Martorell y González, y contó con 54 ítems y cuatro opciones de respuestas (nunca, algunas veces, muchas veces, siempre). El segundo se denomina Cuestionario de Conducta Antisocial, construido por los mismos autores con 36 ítems y con las mismas alternativas. Entre las conclusiones que obtuvieron, podemos mencionar a las siguientes:

En las dos ciudades, los varones resultaron menos empáticos que las mujeres, con más conductas agresivas y menos sentimientos de soledad.

En Buenos Aires, no se observaron diferencias significativas relacionadas con el género en cuanto al aislamiento y agresividad. Las mujeres se percibieron más retraídas que los varones, pero en igual medida que las valencianas.

Los jóvenes de 13 a 15 años (grupo de menor edad), se mostraron más empáticos, sociables y respetuosos que los de 16 a 18 años.



El método CoRT ha sido desarrollado para enseñar a los estudiantes a pensar sistemáticamente. Se basa en diseñar y hacer explícito un instrumento de pensamiento (de Bono, 1994). Se utilizan una serie de tarjetas que son presentadas a los estudiantes y están organizadas en seis secciones que contienen 10 lecciones cada una. Las secciones son las siguientes: amplitud de percepción; organización del pensamiento; interacción, argumentación y pensamiento crítico; pensamiento creativo con herramientas específicas; información y sentimientos y acción. Entre otros objetivos, la primera sección intenta obtener una mayor

amplitud de percepción de ciertos sectores de la realidad. Uno de los instrumentos que utiliza es denominado PNI (abreviatura de positivo, negativo e interesante). Por ejemplo, una lección CoRT podría empezar de la siguiente manera:

“El profesor establecería una tarea simple: por ejemplo, pedirles a los alumnos más jóvenes que hagan un diseño nuevo de la cabeza humana” (de Bono, 1994).

A partir las propuestas de los alumnos, el docente les solicita que mencionen los elementos positivos, negativos e interesantes del diseño propuesto. Los sistematiza de alguna forma y explica que siempre que tengamos que tomar una decisión es conveniente evaluar sistemáticamente los elementos positivos, negativos e interesantes que posee cada alternativa.

Luego, le presenta un problema diferente para que los alumnos trabajen en grupos de cuatro a seis integrantes, solicitando las respuestas ni bien ha finalizado el tiempo asignado y se presentan otras situaciones. Al terminar la lección, se realiza un cierre debatiendo sobre la finalidad y los momentos en que se puede aplicar esta técnica y el docente presenta una síntesis sobre las características de ella.

Edwards (1994) realizó una investigación para evaluar el método CoRT, con 202 alumnos de 7º grado que pertenecían a siete divisiones, tres de una escuela católica y cuatro de una estatal de Australia. En total, 115 alumnos recibieron el tratamiento, conformándose un grupo de control con 87 estudiantes con las otras tres clases. Participaron voluntariamente siete docentes y se los asignó a tratamiento o control con la finalidad de contar con una mayor amplitud de tipos de personalidad entre los docentes.

Cuatro grupos de alumnos fueron distribuidos a tratamiento y recibieron, de su propio maestro, las 10 lecciones CoRT, en aproximadamente 7 a 8 horas de enseñanza. Los otros grupos continuaron con su programa normal. Los docentes recibieron capacitación sobre este método de enseñanza del pensamiento. Tanto los grupos asignados al tratamiento, como los que no lo recibieron, fueron sometidos a pruebas previas, posteriores y diferidas. Estas últimas se realizaron semanas más tarde que las posteriores. Entre los resultados positivos obtenidos, se observaron avances en la aptitud escolar, originalidad y exámenes nor-

males. No se verificaron progresos con relación al autoconcepto como pensador y en la clasificación de los docentes de las técnicas de pensamiento de sus alumnos.



Gualpa (2003) presenta una encuesta implementada a 100 docentes del partido de Chivilcoy. Los datos se procesaron según la antigüedad de los mismos: “con menos de 12 años de antigüedad” y “con más de 12 años de antigüedad”. Entre los resultados hallados, la autora menciona:

“Es visible que el factor de mayor influencia a la hora de elegir la docencia para el grupo de mayor antigüedad fue el geográfico (...) por el contrario entre el grupo de menor antigüedad se presenta una mayor dispersión de las respuestas realizadas (...) Los docentes de menor antigüedad poseen vívida la imagen de simpatía de sus docentes, los de mayor antigüedad destacan otros aspectos entre los que mencionan: la exigencia del docente, la competitividad y la falta de libertad” (:16).



En un estudio (Sposetti y Echevarría, 2004) se indagó si el tiempo dedicado al estudio se relaciona a la lentificación en la universidad. Ésta última se consideró según los alumnos regularizaron o no la materia que cursaron. La primera variable se midió en relación a los días por semana que los alumnos asignaban al estudio, como así también, las horas por día, por un lado, sin obligaciones de exámenes y, por otro lado, ante la posibilidad de rendir algún examen parcial. También se indagó la anticipación con que estudiaron para las evaluaciones.

Se analizó la relación entre las respuestas a cada una de las preguntas que se incluyen a continuación y la lentificación, aunque los datos se recategorizaron del siguiente modo (en el Apéndice de Sposetti de Croatto, Silva de Ducuron y Echevarría, 2004, se hallan las preguntas con las categorías originales):

“Durante el primer año de estudio, cuando NO tenías que rendir un parcial estudiabas” (Tres días o menos por semana, Cuatro o mas días por semana);

“Durante el primer año de estudio, cuando NO tenías que rendir un parcial estudiabas” (Dos o menos horas por día, Tres o mas horas por día);

“Durante el primer año de estudio, por lo general, ¿con cuánta anticipación estudiabas ante un parcial?” (Una semana o menos; Más de una semana);

“Durante el primer año de estudio, por lo general, cuando tenías que rendir un parcial estudiabas” (Todos los días, Cuatro o cinco días, Dos o tres días, Uno o dos días;

“Durante el primer año de estudio, cuando tenías que rendir un parcial estudiabas” (Cinco o más horas por día; Tres o cuatro horas por día; Una o dos horas por día).

Se halló que “la pérdida de la condición de alumno regular se presenta asociada al escaso tiempo dedicado al estudio. La información recogida pone en evidencia que un compromiso activo del estudiante a la hora de planificar su proceso de estudio, conduce al mayor éxito académico, y más específicamente, que para lograrlo con mayor probabilidad, es necesario dedicarle cuatro o más días por semana y tres o más horas por día, cuando no tienen previsto un examen; preparar sus exámenes parciales con más de una semana de anticipación, estudiar por lo menos dos o tres días por semana y dedicarle, en la medida de lo posible, tres o cuatro horas por día” (Sposetti de Croatto y Echevarría, 2004: 58).

Nota: Al analizarlo, tener en cuenta que no necesariamente para cada pregunta se utilizó el mismo diseño.



Roselli et al. (1995) realizaron un trabajo en el que analizaron la relación entre el logro cognitivo individual y el modelo tutorial (tipo de relación que existió entre los alumnos y los docentes). Los modelos

tutoriales que compararon fueron el egocéntrico, asimétrico y cooperativo (o constructivo). Los sujetos pertenecían a escuelas públicas de una ciudad argentina y tenían una edad promedio de 13 años. Se consideraron tres grupos, que

“fueron tres divisiones completas, siendo asignados aleatoriamente a cada una de las condiciones tutoriales.

La cantidad de sujetos de cada grupo fue variable (27, 23, y 25), aunque las características de los tres conglomerados resultaron muy similares” (ob. cit.: 54).

Como instrumento de recolección de datos utilizaron ocho problemas planteados en el Test de Matrices Progresivas de Raven. Éstos se tomaron en forma individual antes y después de la sesión experimental. En esta última se utilizó otra serie de ocho matrices, también tomadas del mismo test, siendo cada uno de ellos equivalente en términos lógicos a los implementados antes y después de la sesión experimental.

En cuanto a los resultados hallados, “la lectura de los datos muestra que, como era esperable, hubo progreso en los tres grupos, siendo más notorio en el que fue afectado a la tutoría constructiva. Este grupo es, precisamente, el de más bajo nivel inicial; los restantes son equivalentes de este aspecto. También el grupo constructivo parece más homogéneo” (ob. cit.: 56).



Afonso y Bueno (2010), realizaron un estudio orientado a indagar si el “tipo de recuerdos autobiográficos evocados a través de un programa individual de reminiscencia” permite reducir los síntomas depresivos en adultos mayores. Según los autores, “se construyó un diseño cuasi-experimental con evaluaciones pretest y posttest del tipo de recuerdos autobiográficos a partir de la presentación de palabras estímulo”. Seleccionaron 90 sujetos mayores de 65 años que presentaban síntomas depresivos y no estaban medicados, y no mostraban signos de demencia. “Los sujetos seleccionados fueron asignados aleatoriamente a las condiciones experimentales: (1) grupo experimental, con sesiones individuales de reminiscencia; (2) grupo control, sin ningún tipo de intervención; (3) grupo control-placebo, con sesiones de relajación”. Los

resultados mostraron, entre otras cosas, en los sujetos asignados a las sesiones de reminiscencia mejoras en relación con los síntomas depresivos, y “aumentos significativos en el número de recuperaciones autobiográficas específicas y positivas (... y) fuertes asociaciones negativas de la sintomatología depresiva con la especificidad de los recuerdos y con su carácter positivo”.



Arredondo et. al. (2004) se propusieron analizar la relación entre la forma de enseñanza y el rendimiento en alumnos de educación básica. La población de interés estuvo integrada por todos los séptimos grados de una Unidad Educativa del Estado Miranda de Venezuela. El contenido se refirió a la Historia de este país. Se eligieron al azar cuatro secciones, asignándose las dos primeras al grupo de control y las dos restantes al experimental. Los pasos que siguieron fueron los siguientes:

“Observación en el aula de clase.

Entrega de encuestas a los profesores y alumnos.

Selección de sujetos.

Asignación de sujetos al grupo control (...) y al experimental (Textocentrismo).

Aplicación del tratamiento (Textocentrismo) a uno de los grupos.

Aplicación del Post-Test a ambos grupos.

Análisis estadístico (t de Student)”.

Una de las conclusiones a que arribaron es expresada en los siguientes términos:

“El uso exclusivo del texto, por parte del docente limita la creatividad en el alumno, no propicia la formación de un individuo activo y crítico, sino por el contrario lo conduce al aprendizaje mecánico o memorístico, sin comprender los contenidos, influyendo negativamente en el rendimiento académico del mismo”.

Bibliografía citada

- Ander-Egg, E. 2000. *Métodos y técnicas de investigación social III. Cómo organizar el trabajo de investigación*. Lumen Hvmanitas. Buenos Aires, Argentina.
- Ander-Egg, E. 2003. *Métodos y técnicas de investigación social IV. Técnicas para la recogida de datos e información*. Lumen Hvmanitas. Buenos Aires, Argentina.
- Ander-Egg, E. y M. Aguilar Idáñez. 2004. *Cómo elaborar un proyecto. Guía para diseñar proyectos sociales y culturales*. Lumen Hvmanitas. Buenos Aires.
- Afonso, R. y Bueno, B. 2010. Reminiscencia con distintos tipos de recuerdos autobiográficos: efectos sobre la reducción de la sintomatología depresiva en la vejez. *Psicothema*. Vol. 22, N° 2, pp. 213-220.
- Arnal, J., D. del Rincón y A. Latorre. 1992. *Investigación educativa*. Fundamentos y metodologías. Labor, Barcelona.
- Arredondo Q., B. Ibelys, R. Zapata y J. Carlos. El Textocentrismo. <http://www.monografias.com/trabajos6/texto/texto.shtml> . Consultado el 21 octubre de 2004.
- Blázquez, G. 2006. Nenas cuarteteras: hegemonía heterosexual y formas de clasificación de las mujeres en los bailes de cuarteto. En Dalmaso, M. y A. Boria (Editoras). *Discurso social y construcción de identidades: mujer y género*. Ediciones del Programa de Discurso Social. Córdoba.
- Bogdan, R. y S. Biklen. 1992. *Qualitative research for education. An introduction and methods*. Allyn and Bacon, Boston.
- Campbell, D. y J. Stanley. 2005. *Diseños experimentales y cuasiexperimentales en la investigación social*. Amorrortu Editores, Buenos Aires.
- Cardozo, G. y A. Alderete. 2009. Adolescentes en riesgo psicosocial y resiliencia. *Psicología desde el Caribe*. N° 23.
- Carreras, M., A. Brizzio, R. González, S. Mele y M. Casullo. 2008. Los estilos de apego en los vínculos románticos y no románticos. Estudio comparativo con adolescentes argentinos y españoles. RIDEP N° 25. Vol. Pp.: 107-124. http://www.aidep.org/03_ridep/R25/R255.pdf. (Consultada: 13-04-2011).
- Castañeda Jiménez, J. 1996. *Métodos de investigación*. McGraw-Hill, México.
- Cea D'Ancona, M. A. 1999. *Metodología cuantitativa: estrategias y técnicas de investigación social*. Síntesis, Madrid.
- Chirinos, J., C. Brindis, V. Salazar, O. Bardales y L. Reátegui. 1999. Perfil de las estudiantes adolescentes sexualmente activas en colegios secundarios de Lima, Perú. *Rev Med Hered* 10 (1).
- Cohen, L. y L. Manion. 1990. *Métodos de investigación educativa*. La Muralla, Madrid, España.
- Creswell, J. 2003. *Research design: qualitative, quantitative, and mixed methods approaches*. Sage Publications, Londres.

- de Bono, E. 1994. La enseñanza del pensamiento en la educación y el método CoRT. En Maclure, S. y P. Davies (comps.) *Aprender a pensar, pensar en aprender*. Gedisa, Barcelona. Pp. 35-51.
- Díez, J. y C. Moulines. 1997. *Fundamentos de Filosofía de la Ciencia*. Ariel: Barcelona.
- Echevarría, H. 2005. *La integración de métodos cualitativos y cuantitativos en la investigación psicogenética y el problema de la validez*. Dunken, Buenos Aires.
- Echevarría, H. 2008a. *La investigación cualitativa y el análisis computarizado de datos*. Homo Sapiens Ediciones. Rosario.
- Echevarría, H. 2008b. *Comprensión lectora de lógica y matemática en alumnos universitarios*. Editorial EFUNARC. Río Cuarto.
- Echevarría, H. 2011a. *Diseño y plan de análisis en investigación cualitativa*. HomoSapiens Ediciones. Rosario.
- Echevarría, H. 2011b. La validez interna de los diseños de investigación. En Wester, J.; G. Müller y L. Martella (Eds.). *Bien común en sociedades democráticas*. En Wester, J.; G. Müller y L. Martella. (Editoras). *Bien común en sociedades democráticas*. Ediciones del ICALA. Río Cuarto.
- Echevarría, H. 2011c. ¿Validez de constructo, validez de los instrumentos de recolección de datos o validez teórica? En Wester, J.; G. Müller y L. Martella. (Editoras). *Bien común en sociedades democráticas*. Ediciones del ICALA. Río Cuarto.
- Edwards, J. 1994. Trabajo de investigación sobre el método CoRT. En Maclure, S. y P. Davies (comps.) *Aprender a pensar, pensar en aprender*. Gedisa. Barcelona.
- Erickson, F. 1989. Métodos cualitativos de investigación sobre la enseñanza. En Wittrock, M. *La investigación de la enseñanza. métodos cualitativos y de observación*. Paidós.
- Gibaja, R. 1988. Acerca del debate metodológico en la investigación educacional. *La Educación*, Nº 103, año XXXII.
- Gil Cuadra, F. y L. Rico. Romero. 2003. Concepciones y creencias del profesorado de secundaria sobre enseñanza y aprendizaje de las matemáticas. *Enseñanza de las Ciencias*. 21 (1), 27-47. <http://www.raco.cat/index.php/ensenanza/article/viewFile/21885/21719>. (Consultada el 20-04-2011).
- González, R., M. Casullo, C. Martorell y A. Calvo. 1998. Evaluación de los comportamientos sociales. Aportaciones de un estudio comparativo. En Casullo, M. *Adolescentes en riesgo. Identificación y orientación psicológica*. Paidós. Barcelona.
- Grillo, M. 2003. *El proyecto de investigación en las Ciencias Sociales. Una propuesta para su elaboración en investigaciones empíricas*. Facultad de Ciencias Humanas. UNRC.
- Gualpa, V. 2003. ¿Ser maestro...? Una mirada reflexiva acerca de la elección por la docencia y la formación recibida. *Aprendizaje Hoy* Nº 54. Año XXIII. Pp. 7-20.
- Hempel, C. 1979. *La explicación científica*. Paidós, Buenos Aires.

- Hernández Pina, F. 1998. Diseños de investigación experimental. En Buendía Eisman, L., P. Colás Bravo y F. Hernández Pina. 1998. *Métodos de investigación en psicopedagogía*. McGraw-Hill. Madrid, España.
- Hernández-Sampieri, R. H., C. Fernández-Collado y P. Baptista. 1991. *Metodología de la investigación*. McGraw-Hill, México.
- Kerlinger, F. 1990. *Investigación del comportamiento*. Mc. Graw-Hill, México.
- León, O. e I. Montero. 1993. *Diseño de investigación*. Mc Graw-Hill, Madrid.
- León, O. e I. Montero. 2003. *Diseño de investigación*. Mc Graw-Hill, Madrid.
- León, O. e I. Montero. 2004. *Métodos de investigación en psicología y educación*. Mc Graw-Hill, Madrid.
- Manheim, H. 1982. *Investigación sociológica*. Ediciones CEAC, Barcelona.
- Maynt, R., K. Holm y P. Hübner. 1975. *Introducción a los métodos de la sociología empírica*. Alianza, Madrid.
- Miles, M. y A. Huberman. 1994. *Qualitative data analysis*. Sage Publications, Londres.
- Miller, W. y B. Crabtree. 1994. Clinical research. En Denzin, N. e Y. Lincoln (Editores). *Handbook of qualitative research*. Sage Publications, California.
- Montiel-Nava, C., J.A. Peña e I. Montiel-Barbero. 2003. Datos epidemiológicos del trastorno por déficit de atención con hiperactividad en una muestra de niños marabinos. *Revista Neurología*. 37 (9). Pp.: 815-819. <http://www.tdahlatinoamerica.org/documentos/articulosnuevos/prevalencia2.pdf> (Consultada el 13-04-2011).
- Padua, J. 1987 *Técnicas de investigación aplicada a las ciencias sociales*. Fondo de Cultura Económica, México.
- Popkewitz, T. 1988. *Paradigma e ideología en la investigación educativa*. Mondadori.
- Roselli, N. 2008. La disyuntiva individual-grupal. Comparación entre dos modelos alternativos de enseñanza en la universidad. *Ciencia, Docencia y Tecnología* N° 36, Pp.: 87-118. <http://www.scielo.org.ar/pdf/cdyt/n36/n36a05.pdf><http://www.scielo.org.ar/pdf/cdyt/n36/n36a05.pdf> (consultada el 09-04-2011).
- Roselli, N.; M. Hechen y L. Gimelli. 1995. Tres modelos de tutoría docente en una tarea de resolución de problemas. *Revista IRICE*, N° 9. Rosario.
- Rusconi, C. 2003. *Algunas convenciones académicas sobre citas y referencias*. En Grillo, 2003.
- Sabino, C. 1996. *El proceso de investigación*. Humanitas. Buenos Aires, Argentina.
- Samaja, J. 1994. *Epistemología y metodología. Elementos para una teoría de la investigación científica*. Eudeba, Buenos Aires.
- Samaja, J. 2004. *Proceso, diseño y proyecto en investigación científica*. Episteme JVE, Buenos Aires.
- Shadish, W., T. Cook y D. Campbell. 2002. *Experimental and Quasi Experimental Designs for Generalized Causal Inference*. Houghton Mifflin, Boston.

- Siegel, S. 1991. *Estadística no paramétrica. Aplicada a las Ciencias de la Conducta*. Trillas: Mexico.
- Sposetti de Croatto, A., N. Silva de Ducuron y H. Echevarría. 2004. *Deserción y lentificación en el pensar y sentir de los alumnos universitarios*. Investigación Evaluativa. Publicado en CD. Río Cuarto, Argentina.
- Sposetti de Croatto, A. y H. Echevarría. 2004. *Tiempo dedicado al estudio y lentificación en la Universidad*. En Sposetti de Croatto, A., N. Silva de Ducuron y H. Echevarría. 2004.
- Tashakkori, A. y Ch. Teddlie. (Editores). 2010. *SAGE handbook of mixed methods social and behavioral research*. Sage Publications. California.
- Teddlie, Ch. y A. Tashakkori. 2009. *Foundations of mixed methods research. Integrating quantitative and qualitative approaches in the social and behavioral sciences*. Sage Publications. California.
- Van Dalen, D. y W. Meyer. 1971. *Manual de técnica de la investigación educacional*. Paidós, Buenos Aires.
- Vidal-Abarca, E. 2000. ¿Son los textos una ayuda para la comprensión? Pp. 141-152. En J. Pozo y C. Monereo (Coord.) *El aprendizaje estratégico. Enseñar a aprender desde el currículo*. Santillana.
- Vuyk, R. 1984. *Panorámica y crítica de la Epistemología Genética de Piaget 1965-1980, I*. Alianza, Madrid.
- Yuni, J. y C. Urbano, 2003. *Técnicas para investigar y formular proyectos de investigación*. Brujas, Córdoba, Argentina.



Los diseños de investigación cuantitativa en psicología y educación

Hugo Darío Echevarría

Se presentan en un lenguaje simple los modelos de investigación cuantitativos posibles de aplicar en Psicología y Educación, con cuadros y esquemas integradores que permiten a los lectores apreciar las diferencias y las interrelaciones que existen entre ellos, se trata el tema de la validez y, fundamentalmente, se esquematizan con misma simbología para evitar las confusiones que crea el uso de distintas convenciones con diferentes diseños. Comienza con una introducción en la que se describe el proceso de investigación, que sirve de marco para una clasificación de los diseños que facilita una visión de conjunto de los mismos. Se presentan los estudios experimentales estrictos, los cuasiexperimentales, los no experimentales, los factoriales y los descriptivos. El libro, trata de reflejar la experiencia del autor de más de 20 años en Cátedras de Metodología de la Investigación y la participación de 30 años en distintos proyectos, como ayudante de investigación, como docente investigador, y como director de proyectos subsidiados por la Secretaría de Ciencia y Técnica de la Universidad Nacional de Río Cuarto y el Instituto de Investigación de la Universidad Nacional de Villa María. Es una guía que sirve de ayuda para seleccionar entre los distintos diseños aquél que puede resultar más adecuado a cada proyecto, explicando en qué consiste cada uno, además de mostrar ejemplos y señalar las ventajas y dificultades que presentan, lo que también se incluye en un cuadro integrador.

